

**Molekularbiologische Untersuchungen zu Funktion und
Phylogenie methanotropher Bakterien**

Dissertation

Zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
(Dr. rer. nat.)

dem
Fachbereich Biologie
der Philipps-Universität Marburg
vorgelegt von

Peter Ricke
aus Arnsberg - Neheim

Marburg a.d. Lahn 2004

Die Untersuchungen zur vorliegenden Arbeit wurden von November 2001 bis Oktober 2004 am Max-Planck-Institut für terrestrische Mikrobiologie, Abteilung Biogeochemie, in Marburg a. d. Lahn unter Leitung von Herrn PD Dr. Werner Liesack durchgeführt.

Vom Fachbereich Biologie der Philipps-Universität Marburg als Dissertation

angenommen am: _____

Erstgutachter: PD Dr. Werner Liesack

Zweitgutachter: Prof. Dr. Wolfgang Buckel

Tag der Disputation: _____

Im Rahmen dieser Promotion wurden Beiträge zu folgenden wissenschaftlichen Publikationen erstellt (in der Reihenfolge des Erscheinens) :

Tchawa Yimga, M., Dunfield, P.F., Ricke, P., Heyer, J. und Liesack, W. (2003). Wide distribution of a novel *pmoA*-like gene copy among type II methanotrophs, and its expression in *Methylocystis* strain SC2. *Applied and Environmental Microbiology* **69**, 5593-602.

Ricke, P., Erkel, C., Kube, M., Reinhardt, R. und Liesack, W. (2004). Comparative Analysis of the Conventional and Novel *pmo* (Particulate Methane Monooxygenase) Operons from *Methylocystis* Strain SC2. *Applied and Environmental Microbiology*. **70**, 3055–3063

Dedysh, N.S., Ricke, P. und Liesack, W. (2004). NifH and NifD phylogenies: an evolutionary basis for understanding nitrogen fixation capabilities of methanotrophic bacteria. *Microbiology*. **150**, 1301–13

Ricke, P., Kolb, S. und Braker, G. Application of a newly developed ARB-integrated *in silico* T-RFLP tool reveals the dominance of a novel *pmoA* cluster in a forest soil. *Applied and Environmental Microbiology*. Im Druck.

Erkel, C., Ricke, P., Kube, M., Reinhardt, R. und Liesack, W. (2004). Retrieval of first genome data for rice cluster I methanogens by a combined approach of cultivation and molecular ecology techniques. *FEMS Microbiology Ecology*. In Revision

Ricke, P., Erkel, C., Kube, M., Reinhardt, R. und Liesack, W. USC α , a group of as-yet uncultured methanotrophs: First insights into the genomic organization substantiates an evolutionary affiliation to the alphaproteobacterial *Methylocapsa acidiphila*. In Vorbereitung

Darüber hinaus sind folgende Veröffentlichungen geplant :

Ricke, P. Kube, M. und Liesack, W. Arbeitstitel: Comparative Analysis of genomic fragments of uncultivated crenarcheota of the cluster 1.1c In Vorbereitung

Erkel, C., Kube, M., Ricke, P., Reinhardt, R. und Liesack, W. Arbeitstitel: Concerned sequencing of the complete genome of an uncultured rice cluster I methanogen. In Vorbereitung

in gedanken an
eine zukunft, die nun bald beginnen sollte,
und die doch nie mehr sein wird,
außer in manchen träumen

Inhaltsverzeichnis

<i>Inhaltsverzeichnis</i>	<i>I</i>
<i>Abkürzungsverzeichnis</i>	<i>IV</i>
<i>Zusammenfassung</i>	<i>V</i>
<i>1 Einleitung</i>	<i>1</i>
1.1 Allgemeine Übersicht	1
1.2 Methan-oxidierende Bakterien	4
1.2.1 Reinkulturen	4
1.2.2 Methanoxidierende Bakterien – Umweltsequenzen	7
1.2.3 Hoch-affine Methanotrophe	9
1.3 Der Kohlenstoff-Stoffwechsel und die pMMO methanotropher Bakterien	12
1.4 Methan-Monooxygenasen	14
1.5 Biochemische Charakterisierung der partikulären Methan-Monooxygenase	16
1.6 Eine neuartige paraloge Genkopie der <i>pmoA</i>	17
1.7 Der Stickstoff-Stoffwechsel methanotropher Bakterien	18
1.8 Ziele der Arbeit	19
<i>2. Material und Methoden</i>	<i>21</i>
2.1 Chemikalien	21
2.2 Sterilisationsverfahren	21
2.3 Verwendete Organismen und Referenzstämme	21
2.4 Anzucht von Mikroorganismen	23
2.4.1 Anzucht methanotropher Bakterien	23
2.4.2 Anzucht von Klonen	23
2.5 Verwendeter Boden	23
2.6 Erstellung von "Large Insert Libraries"	24
2.6.1 "Large Insert Libraries"	24
2.6.2 DNA-Präparation für die Erstellung von BAC-Banken	26
Einbettung in Agarose	26
Zellaufschluß und Aufreinigung der DNA	26
Partielle Restriktion und Größenselektion	27
2.6.3 DNA-Präparation für die Erstellung von Fosmid-Banken	29
DNA-Extraktion aus Waldboden	29
Aufreinigung von Umweltproben	29
Enzymatische Reparatur der Enden gescheerter DNA	30
2.6.4 Größenselektion mittels PFGE	30
2.6.5 Elution größenselektierter DNA aus dem Agarosegel	32
2.6.6 Klonierung mittels BAC-Vektoren	33
Ligation	33
Transformation	34

2.6.7	Klonierung mittels Fosmiden	34
2.6.8	Größenanalyse der BAC- und Fosmid-Inserts von Klonen	35
2.6.9	Vervielfältigung, Screening und Lagerung von LILs	36
	Lagerung und Screening von Mikrotiterplatten	37
	Agarplattenscreening	37
2.7	PCR	38
2.7.1	PCR-Amplifikation	38
2.7.2	Amplifikation von Nitrogenase-Genen aus Reinkulturen	39
2.7.3	Extraktion von DNA aus Reinkulturen für die PCR-Amplifikation	40
2.7.4	Amplifikation von <i>pmoA1</i> und <i>pmoA2</i> aus Reinkulturen und Klonbibliotheken	41
2.7.5	Amplifikation von <i>pmoA</i> -Sequenzen des USC α aus Umweltproben und Klonen	41
2.7.6	PCR basiertes Screening	42
2.8	Analytische Agarose-Gelelektrophorese	42
2.9	Sequenzierungen	43
2.9.1	Sequenzierung von PCR-Produkten	43
2.9.2	Direktsequenzierung ohne vorhergehende PCR -Amplifikation	44
2.9.3	"Shotgun"-Sequenzierung genomischer Fragmente	45
2.10	Bestimmung von Transkriptions-Startpunkten	46
2.10.1	Arbeiten mit RNA	46
2.10.2	Extraktion von Gesamt-RNA mit dem RNeasy (Mini)-Kit	46
2.10.4	Extraktion von Gesamt-RNA mittels Trizol-Reagent	47
2.10.5	Abbau der DNA mittels DNase	48
2.10.6	cDNA-Synthese	48
2.10.7	Bestimmung der Promoter-Regionen mittels RACE	49
2.10.8	Kontrolle der RACE-Ergebnisse	50
2.11	Bioinformatik	51
2.11.1	Datenbanken	51
2.11.2	Phylogenetische Analysen	52
2.11.3	Computergestützte Sekundärstruktur-Analysen von PmoCAB	54
2.11.4	Identifikation konservierter Aminosäure-Reste	55
2.11.5	Signaturanalyse von PmoA1, AmoA und PmoA2	55
2.11.6	Entwicklung spezifischer Primer für die PCR	56
2.11.7	tRFcut	56
2.11.9	ORF-Vorhersage und Gen-Annotation	57
2.11.10	Phylogenetische Zuordnung anhand genomischer Signaturen	58
2.11.11	Weitere verwendete Programme	60
2.11.12	Computergestützte Vorhersage von Promotoren und Terminatoren	61
2.12	Oligonukleotide und Primer	61
3.	Ergebnisse	63
3.1	Untersuchungen von pMMO-2 kodierenden Genen	63
3.1.1	Untersuchungen zu <i>pmoA2</i> , einer neuartigen Genkopie der <i>pmoA</i>	63
3.1.2	Vergleichende Analyse der Operons von pMMO-1 und pMMO-2 von <i>Methylocystis</i> sp. Stamm SC2	64

3.2 Vergleichende Untersuchung von Genom-Fragmenten eines nicht-kultivierten Vertreters des USC α sowie von <i>M. acidiphila</i>	65
3.3 Untersuchungen von <i>nifH</i> und <i>nifD</i> methanotropher Bakterien	66
3.4 tRFcut : <i>in silico</i> tRFLP-Analyse in ARB	67
4. Diskussion	135
4.1 Untersuchungen zu Funktion und Verbreitung der pMMO-2	135
Eine ungewöhnliche paraloge Kopie des <i>pmoA</i> -Gens.	135
4.1.1 Untersuchung zur Verbreitung und Phylogenie der <i>pmoA2</i>	135
4.1.2 Vergleichende Analyse der Operons <i>pmoCAB1</i> und <i>pmoCAB2</i> von <i>Methylocystis</i> sp. Stamm SC2	138
Analyse der abgeleiteten Aminosäuresequenzen von pMMO-1 und pMMO-2	140
Bestimmung der Transkriptionstartpunkte mittels RACE	142
Mögliche Funktion der pMMO-2	143
4.2 Erste Einblicke in ein Genom nicht-kultivierter MB	145
4.2.1 USC α : langgesuchte hoch-affine MB ?	145
4.2.2 Erstellung und Screening von Large-Insert-Bibliotheken	145
4.2.3 Phylogenetische Zuordnung des USC α	147
Analyse konservierter Gene und Gencluster des USC α -Fragments	147
SOM-basierte phylogenetische Analyse der USC α -Genom-Signatur	147
Zuordnung des USC α innerhalb der <i>Alphaproteobacteria</i>	148
4.2.4 Vergleichende Analyse der pMMO-Operons von USC α und <i>M. acidiphila</i>	150
Vergleich von Primär und Sekundärstrukturen	150
Phylogenetische Analyse der pMMO	151
Ein mögliches viertes Gen im pMMO-Operon	152
4.3 Die Nitrogenase-Gene methanotropher Bakterien	155
4.3.1 Einleitung	155
4.3.2 Vergleichende phylogenetische Analyse von NifH und NifD	155
4.3.3 Ungewöhnliche Gruppierung von NifD und NifH von <i>M. capsulatus</i>	156
4.3.4 Phylogenetische Analyse von NifH und NifD acidophiler <i>Alphaproteobacteria</i>	158
4.4 <i>In silico</i> tRFLP-Analyse in ARB mittels des "Tools" tRFcut	160
4.4.1 Einleitung	160
4.4.2 Zuordnung von Restriktionsfragmenten zu entsprechenden phylogenetischen Gruppen	161
4.4.3 Experiment-Design mittels <i>in silico</i> tRFLP-Analyse	162
4.4.4 tRFcut als Alternative zu anderen <i>in silico</i> tRFLP-Programmen	163
4.4.5 Allgemeine Probleme der <i>in silico</i> tRFLP-Analyse	165
4.4.6 tRFcut-spezifische Probleme	166
4.4.7 Anwendung von tRFcut zur Analyse methanotropher Bakterien in pH-neutralen Waldboden	167
Literatur	169
Lebenslauf	183

Abkürzungsverzeichnis

AOB	Ammoniak-Oxidierer
AMO	Ammonium-Monooxygenase
b	basen (kb = Kilobasen, Mb = Megabasen, Gb=Gigabasen)
BAC	Bacterial Artificial Chromosom
BLAST	Basic Local Alignment Search Tool
bp	Basenpaare
DEPC	Diethylpyrocarbonat
DGGE	denaturierende Gradienten-Gelelektrophorese
dNTP	Desoxy-Nukleosidtriphosphat
gdw	gram dry weight
$K_{m(app)}$	apparente halbmaximale Sättigungskonstante
LIL	Large Insert Library
MB	Methanotrophe Bakterien
MMO	Methan-Monooxygenase
ORF	Open Reading Frame
PCR	Polymerase-Kettenreaktion
PLFA	Phospholipid Fatty Acids (Phospholipide)
pMMO	partikuläre Methan-Monooxygenase
PmoA	Untereinheit der partikulären Methan-Monooxygenase
ppmv	parts per million by volume
sp.	Art
spp.	Arten
tRF	terminales Restriktionsfragment
tRFLP	terminaler Restriktionsfragment-Längen-Polymorphismus
upm	Umdrehung pro Minute
USC α	Upland Soil Cluster α
USC γ	Upland Soil Cluster γ
$V_{max(app)}$	apparente maximale Oxidationsrate
v/v	Volumen pro Volumen
w/v	Gewicht pro Volumen

Weitere Abkürzungen wurden entsprechend der „Information for Authors“ des *European Journal of Biochemistry* verwendet.

Zusammenfassung

Die hier beschriebenen Studien dienten der Charakterisierung von methanotrophen Bakterien (MB) mittels molekular- und mikrobiologischer Techniken sowie mittels Methoden der angewandten Bioinformatik. Die Anwendung dieses Methoden-Bestocks auf verschiedene konkrete Fragestellungen schlägt sich im kumulativen Charakter dieser Arbeit nieder. Vorgestellt werden Arbeiten zu folgenden, voneinander abgegrenzt dargestellten Themen: i) Untersuchung der Verbreitung der pMMO-2, einer neuartigen paralogen Kopie der partikulären Methan-Monooxygenase (pMMO-1), eines Schlüsselenzyms der MB sowie die vergleichende Analyse der Operonstrukturen der für pMMO-1 bzw. pMMO-2 kodierenden Gene, ii) Charakterisierung der potenziell zu Methan hoch-affinen MB des "Upland Soil Cluster α " (USC α) mit Methoden der Metagenomik, iii) vergleichende Analyse der für Enzyme der N₂-Fixierung kodierenden Gene *nifH* und *nifD* methanotropher Bakterien sowie iv) Entwicklung eines Computerprogramms für in silico tRFLP-Analysen sowie dessen Erprobung im Rahmen der populationsökologischen Charakterisierung einer MB-Population.

i) Die kultivierten MB wurden aufgrund phylogenetischer, morphologischer und physiologischer Eigenschaften in Typ I, Typ II und Typ X MB eingeteilt. Anhand des für eine Untereinheit der pMMO-2 kodierenden *pmoA2*-Gens konnte die weite Verbreitung dieses bislang uncharakterisierten Enzyms innerhalb der Typ II MB dargestellt werden. Auffällig ist, daß manche MB *pmoA2* nicht enthalten, obwohl das Gen für phylogenetisch wie physiologisch nah verwandte Organismen nachgewiesen wurde. *pmoA2* wurde in keinem der untersuchten Typ I MB oder Typ X gefunden.

Die Erstellung einer BAC-Bibliothek erlaubte die vergleichende Sequenzanalyse der für pMMO-1 bzw. pMMO-2 kodierenden Gene des Referenz-Organismus *Methylocystis* sp. Stamm SC2. Beide Operons bestehen aus drei Genen, die in der Reihenfolge *pmoCAB* organisiert sind. Im Zuge dieser Arbeit wurde die Anzahl bekannter *pmo*-Operons mehr als verdoppelt (vgl. Punkt ii). Die zuvor für Typ II und Typ X MB bekannte Gen-Anordnung ist somit bei allen bisher untersuchten MB hochkonserviert.

Trotz deutlicher Sequenz-Unterschiede der kodierenden Gene ähnelt pMMO-2 sowohl hinsichtlich der Operon-Struktur als auch der abgeleiteten Aminosäure-

Sequenz stark der pMMO-1. Vergleichende Analysen der abgeleiteten Sekundärstrukturen sowie hochkonservierter Aminosäuren führten dazu, daß der pMMO-2 die Funktion einer Methan-Monooxygenase zugeordnet werden konnte. Mittels der 5'-RACE-Technik wurde die Transkription der pMMO-2 nachgewiesen und die Promotoren beider Operons identifiziert.

ii) Obwohl die biogene Oxidation atmosphärischen Methans nachgewiesen ist, wurden bis heute keine MB isoliert, die eine entsprechend hohe Affinität zu Methan aufweisen und die zu Wachstum unter atmosphärischen Methan-Konzentrationen befähigt sind. Die methanotrophe Population von als biogene Methan-Senken charakterisierten, aciden "Upland"-Böden wird durch Vertreter des USC α dominiert. Im Rahmen dieser Arbeit wurde erstmals ein größeres genomisches Fragment eines Repräsentanten dieser bislang nicht kultivierten MB vergleichend analysiert. Dazu wurden die Methoden der Metagenomik etabliert und eine ca. 250.000 Klone umfassende Fosmid-Bibliothek aus Waldboden-DNA erzeugt. Mittels PCR-basierten Hochdurchsatz-Screenings konnten zwei Klone mit USC α -spezifischen Inserts identifiziert werden. Dabei diente der zentrale Bereich der *pmoA* als Markergen. In phylogenetischen Verrechnungen gruppieren partielle *pmoA*-Sequenzen des USC α gemeinsam mit dem *pmoA*-Gen des einzigen kultivierten Vertreters der Gattung *Methylocapsa*, *M. acidiphila*. Deshalb wurde eine Genom-Bibliothek von *M. acidiphila* konstruiert und ein 100 kb großes Fragment zu Referenzzwecken sequenziert. Die vergleichende Analyse der erzielten Sequenzen gegen ca. 200 vollständig sequenzierte prokaryotische Genome bestätigte die nahe Verwandtschaft von USC α und *M. acidiphila* und ermöglichte die eindeutige Zuordnung der Mitglieder des USC α zu den *Alphaproteobacteria*.

Die Analyse der für die pMMO kodierenden Gene ergab für beide Genom-Fragmente die Anordnung in einem Operon in der Reihenfolge *pmoCAB*. Die vergleichende Analyse der Primär- und Sekundärstruktur der abgeleiteten Aminosäure-Sequenzen ergab keine Hinweise auf eine erhöhte Affinität der pMMO des USC α zu Methan, sondern hohe Übereinstimmungen mit den abgeleiteten Strukturen anderer pMMOs. Allerdings konnte stromabwärts des *pmo*-Operons ein Gen identifiziert werden, das sowohl bei *M. trichosporium* Stamm OB3b als auch bei den homologen AMO-Operons nitrifizierender Bakterien an gleicher Stelle lokalisiert ist und möglicherweise ein viertes Gen des *pmo*-Operons darstellt.

iii) Die Erhebung von 25 *nifH*- und 15 *nifD*-Sequenzen ermöglichte erstmals eine umfassende vergleichende Analyse dieser Gene methanotropher Bakterien und damit einhergehend die Zuordnung von in kultivierungsunabhängigen Studien erzielten *nif*-Sequenzen zur Gruppe methanotropher Bakterien. Die Phylogenien der von *nifH* und *nifD* abgeleiteten Aminosäure-Sequenzen stehen in hoher Übereinstimmung mit der Phylogenie der 16S rRNA-Gene. Die einzigen Ausnahmen bilden die Positionen von *M. capsulatus* sowie die unerwartet hohe Sequenz-Ähnlichkeiten von *M. acidiphila* und heterotrophen *Alphaproteobacteria* der Gattung *Beijerinckia*. Die molekularen Daten deuten auf einen alphaproteobakteriellen Ursprung des *nif*-Genclusters von *M. capsulatus* (*Gammaproteobacteria*) hin und implizieren die Übertragung dieser Gene durch horizontalen Gentransfer. Dies korrespondiert mit den physiologischen Daten, da *M. capsulatus* im Gegensatz zu allen anderen gammaproteobakteriellen MB wie auch die alphaproteobakteriellen Typ II MB die Fähigkeit zur N₂-Fixierung unter hohen O₂-Konzentrationen besitzt. Ebenso korreliert die hohe Sequenz-Ähnlichkeit der *nif*-Gene von *M. acidiphila* und *Beijerinckia* spp. mit biochemischen Übereinstimmungen. Die für die klassischen Diazotrophen *Beijerinckia* spp. charakteristische N₂-Fixierung unter atmosphärischen O₂-Konzentrationen wird auch von *M. acidiphila*, jedoch von keinem anderen MB gezeigt.

iv) Ein weiter Gegenstand dieser Arbeit war die Erweiterung des Software-Pakets ARB um ein Werkzeug für die *in silico* tRFLP-Analyse molekularer Sequenzdaten. tRFcut ermöglicht die Zuordnung experimentell erzielter tRFLP-Daten ebenso wie die pre-experimentelle Optimierung der experimentellen tRFLP-Parameter. Im Gegensatz zu den bisher verfügbaren Programmen ist tRFcut hinsichtlich der Auswahl an Primern und Restriktionsenzymen nahezu nicht limitiert. Während herkömmliche tRFLP-Programme auf die Analyse von 16S rDNA-Sequenzen beschränkt sind, können mit tRFcut auch beliebige Datensätze funktioneller Gene (beispielsweise *pmoA*, *nifH*, *nifD*) analysiert werden. Mittels tRFcut wurden die tRFLP-Muster der methanotrophen Population eines pH-neutralen Waldbodens untersucht. Es konnte gezeigt werden, daß Methanotrophe des bislang wenig untersuchten Cluster 1 die dominierenden MB dieses Habitats sind.

1 Einleitung

Im Folgenden wird zunächst eine einführende Übersicht über das Spurengas Methan, seine Quellen und Senken, sowie über die an seiner Umsetzung beteiligten Organismen und Enzyme gegeben. Die für diese Arbeit relevanten Themen werden in den daran anschließenden Unterkapiteln weiter ausgeführt.

1.1 Allgemeine Übersicht

Atmosphärisches Methan gilt, nach Kohlendioxid, als das Treibhausgas mit dem größten Einfluß auf das Klima der Erde. Schätzungen gehen davon aus, daß ca. 20% der unter dem Begriff „Klimawandel“ zusammengefassten Effekte auf eine Veränderung der Konzentration dieses Gas in der Atmosphäre zurück zu führen sind. Während die Konzentration atmosphärischen Methans während der letzten 160.000 Jahre nahezu unverändert blieb (Cicerone & Oremland, 1988), mußte für die vergangenen 200 Jahre ein Anstieg von ca. 0,7 ppmv auf 1,75 ppmv verzeichnet werden. Zur Zeit beträgt der durchschnittliche Anstieg der Methan-Konzentration in der Atmosphäre ca. 0,005 ppmv pro Jahr (gemittelter Wert der Jahre 1990 bis 2002, Cunnold et al. 2002).

Der globale Methankreislauf mit den maßgeblichen Prozessen der Oxidation von Methan und der Methanogenese wird sowohl durch biotische wie auch durch abiotische Prozesse bestimmt. Während der Methanogenese wird Methan von methanogenen *Archaea* entweder als Endprodukt des anaeroben Abbaus organischer C-Verbindungen oder durch Reduktion von CO₂ gebildet. Die Methanogenese kann nach den umgesetzten Edukten (Acetat, H₂/CO₂ bzw. methylierte C₁-Komponenten) in acetoklastische, CO₂-reduzierende und methylotrophe Methanogenese unterteilt werden. Rund 80% des Methans entstehen durch biologische Prozesse (Heyer et al., 2002b). Die wichtigsten biogenen Quellen für atmosphärisches Methan sind natürliche und anthropogene Feuchtgebiete wie Moore und Reisfelder (Conrad, 1996) sowie die Verdauungstrakte von Wiederkäuern und Termiten (Crutzen, 1991), wo Methan bei der Vergärung der Futtermasse durch Mikroorganismen entsteht. Unter den anthropogenen, nicht biogenen Quellen sind beispielsweise der Kohlebergbau und

die Petrochemische Industrie sowie die Biomasseverbrennung von Relevanz (Wuebbles and Hayhoe, 2001). Auch die natürliche Freisetzung von fossil oder in Form von Erdgas gebundenem Methan, beispielsweise entlang von Rissen in der geologischen Formation des Rheintalgrabens (Thielemann, 2000), trägt zum globalen Methanhaushalt bei.

Den zahlreichen unterschiedlichen Quellen stehen wenige Kategorien von Methan-Senken gegenüber: rund 90% des umgesetzten atmosphärischen Methans werden durch photochemische Reaktionen mit Hydroxylradikalen in der Troposphäre abgebaut, ca. 5% gelangen durch Diffusion in die Stratosphäre. Die einzige bekannte biologische Senke stellt die aerobe Oxidation von Methan zu CO₂ durch methanotrophe Bakterien (MB) dar (Wuebbles et al., 1999). Kürzlich wurde eine weitere biogene Form der Methan-Oxidation beschrieben. Die von Konsortien methanotropher *Archaea* und sulfatreduzierender *Bacteria* katalysierte anaerobe Methan-Oxidation dient jedoch nicht dem Abbau atmosphärischen Methans. Die biologischen Prozesse katalysieren im Vergleich zur photochemischen Reaktion nur einen relativ geringen Teil der Oxidation *atmosphärischen* Methans. Dennoch spielt die biogene Methanoxidation eine nicht zu unterschätzende Rolle im globalen Methanhaushalt, da sie vornehmlich an den Methan-Quellen auftritt, wo Methan in hohen Konzentrationen vorliegt und somit die Methan-Netto-Emission drastisch reduziert. Bis 95% des anaerob produzierten Methans werden während der Passage durch die angrenzenden oxischen Schichten, beispielsweise an der Bodenoberfläche oder im wurzelumgebenden Boden oxidiert, bevor sie in die Atmosphäre gelangen können (Conrad und Rothfuss, 1991; Banker et al. 1995; Gilbert und Frenzel, 1995; Denier van der Gon und Neue, 1996; Lombardi et al., 1997; Bosse und Frenzel, 1998; Krüger et al., 2000).

Als maßgebliche Ursache für den Anstieg der atmosphärischen Methankonzentration wird das globale Bevölkerungswachstum gesehen. Mit der Zunahme der Bevölkerungsgröße korreliert einerseits ein Anstieg der Zahl der Methan-Quellen, beispielsweise durch die Zunahme der Rinderbestände, die verstärkte Biogasproduktion, eine erhöhte Anzahl an Mülldeponien und den ausgedehnten und intensivierten Reisanbau. Andererseits werden biogene Methan-Senken z.B. durch Rodungen von Waldflächen oder Überdüngung und mechanische Beanspruchung von Bodenflächen vernichtet.

Als Methan-Senken fungieren hauptsächlich terrestrische Ökosysteme, insbesondere oxische Böden wie Wald- und Wiesenböden (sog. „Upland soils“). Methanaufnahme wurde für Böden der gemäßigten, subtropischen und tropischen Klimazone nachgewiesen. Harris und Mitarbeiter konnten 1982 erstmals zeigen, daß zeitweilig ausgetrocknete Sumpfböden als Senke atmosphärischen Methans fungieren können (Harris et al., 1982). Diese Eigenschaft konnte später auch für tropische Böden (Seiler, 1984), Torfböden (Yavitt et al., 1990a), Tundraböden (Whalen und Reeburgh, 1990), Wald- und Wiesenböden der gemäßigten Zonen (Born et al., 1990; Steudler et al., 1989) und bebaute Ackerböden (Hütsch et al., 1993) nachgewiesen werden. Die jährliche globale Methanaufnahme solcher Böden liegt Schätzungen zufolge zwischen 20 Tg y⁻¹ und 100 Tg y⁻¹ (Dörr et al., 1993; Roslev et al., 1997; Milich, 1998) was einem prozentualen Anteil an der Brutto-Methanabnahme von ungefähr 5% entspräche.

Methan wird durch Methan-Monooxygenasen (MMO) zu Methanol oxidiert, wobei molekularer Sauerstoff zu Wasser reduziert wird. Bei der anschließenden Oxidation des Methanols zu CO₂ entstehen Formaldehyd und Formiat als Stoffwechsel-Intermediate. Die Kohlenstoff-Assimilation erfolgt ausgehend von diesen Zwischenprodukten. Das Schlüsselenzym der Methan-Oxidation, die MMO kommt in zwei phylogenetisch nicht miteinander verwandten Formen vor. Die lösliche MMO (engl.: soluble Methane Monooxygenase, sMMO) liegt frei im Cytoplasma der Zelle vor. Sie ist in vielen, aber nicht allen MB vorhanden. Wesentlich weiter verbreitet ist die membrangebundene, partikuläre MMO (pMMO) (vgl. 1.2.2; 1.3 - 1.5). Die pMMO ist mit einem intrazellulären Membransystem (IZM-System, Abb.1.1) assoziiert, das in verschiedenen MB-Gruppen eine jeweils charakteristische Morphologie aufweist. Die Hydroxylase-Untereinheit der pMMO besteht aus drei Polypeptiden, die durch drei konsekutiv in einem Operon organisierte Gene (*pmoA*, *pmoB* und *pmoC*) kodiert werden. Aufgrund der nahezu universellen Verbreitung der pMMO innerhalb methanotropher Bakterien wird der zentrale Bereich des *pmoA*-Gens häufig als Markergen in Umweltstudien genutzt (vgl. 1.2.2). Ähnliche, wenngleich nicht so intensive Verwendung fanden auch die Gene *mmoX* und *mxoF* (Horz et al., 2001). *mmoX* kodiert für das α-Polypeptid der Hydroxylase-Untereinheit der sMMO (Murrell und Radajewski, 2000), *mxoF* für Methanol-Dehydrogenase (vgl. 1.3).

1.2 Methan-oxidierende Bakterien

1.2.1 Reinkulturen

Die Fähigkeit, Methan zu oxidieren, ist innerhalb der Domäne *Bacteria* auf wenige Gruppen beschränkt, die zu den Phyla *Alpha*-, *Beta*- und *Gammaproteobacteria* gehören. Als MB werden in der Literatur allerdings ausschließlich solche Arten bezeichnet, die mit Methan als einziger Energie- und Kohlenstoff-Quelle in Gegenwart molekularen Sauerstoffs wachsen können. Alle bekannten MB gehören entweder zu den *Alpha*- oder zu den *Gammaproteobacteria*. Davon getrennt zu sehen sind die Ammoniak-oxidierenden Bakterien (AOB, *Beta*- und *Gammaproteobacteria*), die zwar in einer Nebenreaktion ebenfalls Methan verstoffwechseln können, hauptsächlich jedoch CO₂ als Kohlenstoff- und Ammoniak als Energiequelle nutzen. Viele MB können neben Methan auch Methanol, ein Zwischenprodukt des oxidativen Methanabbaus, als Substrat nutzen (Whittenbury *et al.* 1970). Wenige Arten nutzen darüberhinaus auch Methylamin, Formiat oder Formamid (Bowman *et al.*, 1993).

Aufgrund morphologischer, physiologischer und biochemischer Merkmale teilten Whittenbury und Mitarbeiter (1970) die MB in Typ I und Typ II Methanotrophe ein. Die Anordnung des intrazellulären Membransystems diene als wichtigstes morphologisches Unterscheidungsmerkmal. Im Gegensatz zu den in Stapeln angeordneten IZM der Typ I MB, sind die IZM von Typ II MB peripher angeordnet und verlaufen parallel zur äußeren Zellmembran (vgl. Abb. 1.1). Basierend auf der phylogenetischen Analyse ihrer 16S rRNA-Gene gehören Typ I MB zur Gruppe der *Gammaproteobacteria*. Sie besitzen als charakteristische Membranlipid-Fettsäuren (Phospholipid-Fettsäuren, PLFAs) 16:1 ω 7c und 16:0 Fettsäuren. Typ II MB hingegen gehören zu den *Alphaproteobacteria*. Ihre Fettsäuremuster werden von 18:1 Fettsäuren dominiert. Typ I MB verfügen über die Enzyme des Ribulose-Monophosphat-Weges der Kohlenstoff-Assimilation, während Typ II MB Kohlenstoff über den Serin-Weg fixieren. Ausgangspunkt beider Wege ist

Formaldehyd, das als Zwischenprodukt der Methanoxidation entsteht. Ein dritter, von Formiat ausgehender Weg wurde für *Methylococcus capsulatus* (*Gammaproteobacteria*) nachgewiesen: der Benson-Calvin-Zyklus. Mittels dieses Stoffwechselweges vermag *M. capsulatus* im Gegensatz zu allen anderen bekannten MB auch CO₂ zu fixieren. Aufgrund dieser und anderer Besonderheiten wurde *M. capsulatus* im Typ-Konzept häufig nicht als Typ I, sondern als Typ X MB klassifiziert.

Bowman und Mitarbeiter nahmen 1995 eine Neuklassifikation der MB vor, in deren Verlauf die Typ I und Typ X MB zur Familie *Methylococcaceae* zusammengefaßt und die Typ II MB als *Methylocystaceae* angesprochen wurden. Neben den zu diesem Zeitpunkt bekannten Gattungen *Methylococcus*, *Methylobacter*, *Methylobacter* (*Methylococcaceae*) sowie *Methylosinus* und *Methylocystis* (*Methylocystaceae*) sind zwischenzeitlich sechs weitere Gattungen beschrieben worden. *Methylocaldum* (Bodrossy et al., 1997), *Methylosphaera* (Bowman et al., 1997), *Methylosarcina* (Wise et al., 2001) und *Methylohalobium* (Heyer et al., 2004) wurden aufgrund physiologischer, morphologischer und phylogenetischer Eigenschaften den *Methylococcaceae* zugeordnet. Zwei weitere neu beschriebene Gattungen, *Methylocella* (Dedysh et al., 2000) und *Methylocapsa* (Dedysh et al., 2002), konnten keiner der bisher beschriebenen Familien zugeordnet werden. Die bislang isolierten Spezies der Gattung *Methylocella*, *M. sylvestris*, *M. tundrae* und *M. palustris*, besitzen im Gegensatz zu allen Typ I, Typ II und Typ X MB keine pMMO. Das IZM ist nicht oder nur sehr rudimentär vorhanden. Für die Gattung *Methylocapsa* wurde bislang eine einzige Spezies, *M. acidiphila*, beschrieben. *M. acidiphila* verfügt über ein ungewöhnliches intrazelluläres Membransystem. Die IZM sind wie bei Typ I und Typ X MB in Stapeln angeordnet, allerdings sind sie auf einen Teil der Zelle beschränkt (Abb.1.1). Diese neuartige IZM-Anordnung wurde als Typ III IZM-System bezeichnet (Dedysh et al., 2002). *Methylocapsa* und *Methylocella* besitzen 18:1 ω7c als Hauptfettsäure, sodaß sich auch das Fettsäureprofil beider Gattungen von dem der Typ II MB unterscheidet. Sowohl *Methylocapsa* wie auch *Methylocella* sind auf Basis ihrer 16S rRNA-Gene phylogenetisch näher mit den acidophilen heterotrophen Bakterien der Gattung *Beijerinckia* (*Alphaproteobacteria*) als mit anderen MB verwandt. Aufgrund der

phylogenetischen Position wurden beide Gattungen der Familie *Beijerinckiaceae* zugeordnet.

Die klassischen Typ I und Typ II MB dominieren die methanotrophen Gemeinschaften in dauerhaft und regelmäßig gefluteten pH-neutralen Böden wie Reisfeldern oder Mooren, wo sie an den oxisch-anoxischen Grenzschichten die Methan-Netto-Emission dieser Böden verringern. Dagegen sind *Methylocella* spp. und *Methylocapsa* spp. die dominierenden MB in vielen sauren Feuchtgebieten, beispielsweise in *Sphagnum*-besiedelten Hoch- und Zwischenmooren (Dedysh et al., 1998a 1998b und 2000).

Kürzlich isolierten Pacheco-Oliver und Mitarbeiter neuartige MB aus Tundra-Böden (Pacheco-Oliver et al., 2002). Obwohl diese MB zu den *Alphaproteobacteria* gehören, bilden ihre *pmoA*-Sequenzen eine eigenständige, von allen anderen *pmoA* deutlich getrennte Entwicklungslinie ("Cluster I", vgl. 3.5). Verbreitung und Taxonomie dieser Gruppe sind bislang kaum erforscht.

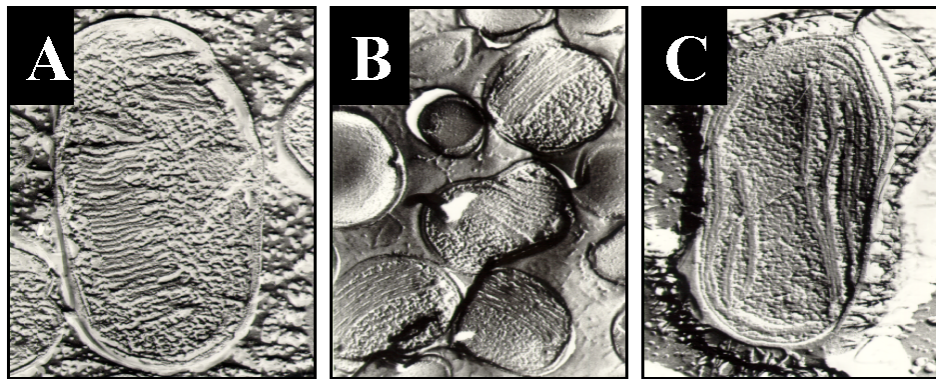


Abb. 1.1 Intrazelluläres Membran (IZM)-System von A) *Methylobacter bovis* 89 (Typ I), B) *M. acidiphila* und C) *M. trichosporium* 19 (Typ II). Die IZM von Typ I MB sind in Stapeln angeordnet, die IZM von Typ II MB hingegen periphär und parallel zur Zellmembran. Die IZM von *M. acidiphila* ist auf einen Teil der Zelle beschränkt und unterscheidet sich deutlich von den bekannten IZM. Sie wurde deshalb als Typ III IZM bezeichnet. Abbildungen modifiziert nach Dedysh et al., (2002).

1.2.2 Methanoxidierende Bakterien – Umweltsequenzen

Für Studien der Abundanz und Verbreitung methanotropher Populationen, aber auch für die Erstellung 16S rDNA-unabhängiger Phylogenien, wurde in den letzten Jahren vermehrt ein Abschnitt des *pmoA*-Gens als molekularer Marker genutzt. *pmoA* kodiert für die β -Untereinheit der pMMO (vgl. 1.4). Aufgrund der nahezu universellen Verbreitung der pMMO, der hohen Konserviertheit ihrer Sequenzen sowie der Tatsache, daß bisher kein horizontaler Gentransfer für dieses Gen nachgewiesen werden konnte, eignet sich *pmoA* ausgezeichnet als funktionelles Markergen. Im Gegensatz zur Verwendung von 16S rRNA-Genen für das Monitoring Methanoxidierender Populationen ermöglicht die Verwendung von *pmoA* die Diskriminierung aller Sequenzen nicht-methanotropher Organismen (mit Ausnahme des *amoA*-Gens nitrifizierender Bakterien, vgl. 1.4). Aufgrund der phylogenetischen Zugehörigkeit Methanotropher sowohl zu den *Alpha*- wie *Gammaproteobacteria* ist es nicht möglich, Primer zu formulieren, die eine Amplifikation der 16S rRNA-Gene aller MB bei gleichzeitiger Diskriminierung nicht-methanotropher *Bacteria* gewährleisten würden. Dagegen wurden verschiedene PCR-Assays beschrieben, welche die *pmoA*-Gene aller bekannten MB erfassen (Holmes et al., 1995, Kolb et al., 2003). Diese universellen Primer detektieren neben *pmoA* lediglich das homologe Gen *amoA*. Die vergleichende Analyse der *pmoA/amoA*-Sequenzen erlaubt eine zur 16S rDNA-basierten Phylogenie bis auf die Spezies-Ebene koherente phylogenetische Zuordnung (Abb. 1.2). Die Analyse funktioneller Markergene ermöglicht also neben der Identifikation bekannter Taxa und der phylogenetischen Zuordnung unbekannter Organismen auch deren gleichzeitige Zuordnung zu der untersuchten funktionalen Gruppe (Gilde), und damit verbundenen die Auflösung von Struktur und Zusammensetzung dieser Gilde.

Letzterem Punkt ist in den vergangenen Jahren eine besondere Bedeutung zugekommen. Mit der Verbreitung molekularer Methoden, insbesondere der Verwendung von *pmoA* als Markergen, gelang vermehrt der Nachweis, daß die in Reinkulturen vorhandenen, gut untersuchten MB in manchen Habitaten nur eine untergeordnete Rolle spielen (Kolb et al., 2003; Kolb, 2004; Knief et al., 2003b).

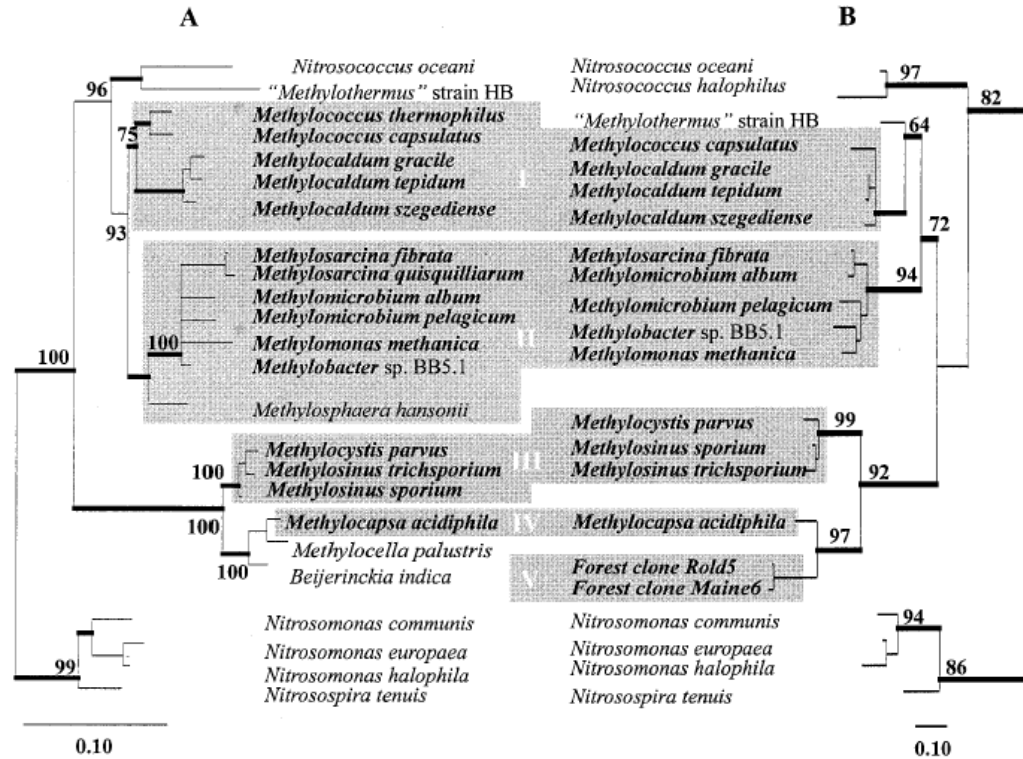


Abb. 1.2: Vergleich von A) 16S rDNA-basierten und B) *pmoA/amoA*-basierten Phylogenien. Beide Bäume wurden mittels Maximum-Likelihood-Methoden (TreePuzzle) berechnet. Stark hervorgehobene Äste wurden durch ProtML-Analysen bestätigt. (I: Typ X MB, II: Typ I MB, III: Typ II MB; IV: *Methylocapsa*, V: Umweltsequenzen des "Upland Soil Cluster α " (USC α). [Aus : Kolb et al., 2003].

Beispielsweise konnten in den letzten fünf Jahren in sogenannten "Upland-Soils", also in überwiegend nicht gefluteten Böden, *pmoA*-Sequenzen nachgewiesen werden, die keiner der bekannten Gruppen methanotropher Bakterien zuzuordnen waren. Sowohl Holmes und Mitarbeiter (Holmes et al., 1999) als auch Henckel und Mitarbeiter (Henckel et al., 2000a) erhielten Sequenzen, die in vergleichenden Analysen Ähnlichkeiten zu *pmoA*-Sequenzen der *Methylocystaceae* aufweisen, jedoch in phylogenetischen Berechnungen eindeutig getrennt verzweigen. Aufgrund der Fundorte, saurer Waldböden, wurde die Sequenzgruppe zunächst als „Forest-Soil Cluster“ angesprochen (Holmes et al., 1999). Jüngere Studien von Knief und Mitarbeitern (2003b) wiesen jedoch eine weitere Verbreitung der Gruppe nach, weshalb eine Umbenennung in "Upland Soil Cluster" erfolgte. Sowohl die Position der *pmoA*-Sequenzen in phylogenetischen Berechnungen wie auch die Ergebnisse von Analysen charakteristischer PFLAs (vgl. 1.2.3) deuten darauf hin, daß die

unbekannten MB phylogenetisch zu den *Alphaproteobacteria* gehören könnten. Dafür spricht auch die auf Basis von *pmoA*-Sequenzen postulierte Verwandtschaft zu *Methylocapsa acidiphila* (Vgl. Abb. 1.3). Knief und Mitarbeiter benannten diese Gruppe entsprechend in " Upland Soil Cluster α " (USC α) um.

Eine weitere Gruppe ungewöhnlicher *pmoA*-Sequenzen wurde in pH-neutralen bis alkalischen (pH > 6,0) Upland-Böden nachgewiesen (Knief et al., 2003b). Die Sequenzen weisen hohe Ähnlichkeiten zu den *pmoA*-Sequenzen der *Methylococcaceae* auf. Untersuchungen dieser Böden durch Markierungsstudien mit $^{13}\text{CH}_4$ resultierten im Nachweis von für *Gammaproteobacteria* charakteristischen Fettsäuren. Dementsprechend wurde das Cluster „USC γ “ benannt. (Knief et al., 2003b; Knief, 2004). Die Verbreitung und Abundanz beider Gruppen wurde gut untersucht. So konnten Kolb und Mitarbeiter beispielsweise für einen tonigen, pH-neutralen (pH 7,7) Waldboden USC γ -Sequenzen und für einen sauren (pH 4,3), sandigen Boden USC α -Sequenzen mittels Real-Time-PCR quantifizieren (Kolb et al., 2003). Beide Gruppen dominierten im jeweiligen Habitat die methanotrophen Lebensgemeinschaften mit einem nachgewiesenen Anteil von mindestens 85%. Von besonderer Bedeutung sind die beschriebenen Umweltsequenzen, da die Standorte, an denen diese neuen Gruppen dominant auftreten, als Senken atmosphärischen Methans beschrieben sind. Bislang war es nicht möglich, diejenigen MB zweifelsfrei zu identifizieren, welche die Methan-Oxidation in diesen Böden katalysieren (vgl. 1.2.3).

Weitere, bisher lediglich in Umweltproben nachgewiesene Sequenzen gruppieren in phylogenetischen Verrechnungen zwischen den Sequenzen der homologen Gene *pmoA* und *amoA* (Henckel et al., 2000b; Holmes et al., 1999b; Reay 2001), sodaß eine eindeutige Charakterisierung der zugehörigen Organismen als MB oder Nitrifizierer zur Zeit nicht möglich ist.

1.2.3 Hoch-affine Methanotrophe

Bereits 1992 berechneten Bender und Conrad für Waldböden nach Vorinkubation mit Methan zwei von der Methan-Konzentration des Bodens abhängige, voneinander stark abweichende Michaelis-Menten-Kinetiken (Bender und Conrad, 1992). Sie unterschieden eine hoch-affine Kinetik mit apparenter halbmaximaler Substrataffinität ($K_{M(\text{app})}$) im nM Bereich von einer niedrig-affinen Kinetik mit $K_{M(\text{app})}$ -Werten im auch

von MB-Reinkulturen bekannten μM Bereich. Die hoch-affine Kinetik wurde bei Methan-Konzentrationen (1,8 ppmv bis 5 ppmv) gemessen, die ungefähr der atmosphärischen Konzentration entsprechen. Für Upland-Böden wurden $K_{\text{M(app)}}$ -Werte im Bereich von 1 – 280 nM sowie apparente maximale Oxidationsraten ($V_{\text{max(app)}}$) zwischen 0,5 und 150 $\text{nmol h}^{-1} \text{gdw}^{-1}$ gemessen (Bender und Conrad, 1992; Bender und Conrad, 1993; Dunfield und Knowles 1995, Whalen und Reeburgh, 1996). Zum Vergleich: für Habitate mit hohen Methan-Konzentrationen werden $K_{\text{M(app)}}$ -Werte im Bereich von 1,7 – 11 μM angegeben (Bender und Conrad, 1993; Dunfield and Knowles 1995, Whalen und Reeburgh, 1996). Inkubation der Upland-Böden mit Methan in Konzentrationen >1000 ppmv führte zum Auftreten der niedrig-affinen Aktivität (Bender und Conrad, 1995; Schnell and King, 1995). Es wurde vermutet, daß der Effekt durch eine von hohen Methankonzentrationen induzierte Aktivität niedrig-affiner MB hervorgerufen wird, und das sowohl niedrig-affine MB als auch bislang nicht kultivierte hoch-affine MB in den Böden vorkommen und in Abhängigkeit von der Substratversorgung aktiviert werden. Dafür spricht, daß einerseits Vertreter der Gattungen *Methylobacterium*, *Methylobacter*, *Methylobacterium*, *Methylobacter* und *Methylobacter* aus „Upland“-Böden isoliert bzw. angereichert werden konnten (Dunfield et al., 1999; Dunfield et al., 2003; Jensen et al., 1998, Fjellbirkeland, 2001), molekulare Analysen von *pmoA*-Sequenzen und Fettsäureprofile andererseits die Dominanz bislang nicht charakterisierter methanotropher Gruppen (USC α und USC γ) belegen.

Das von verschiedenen Autoren postulierte Konzept nicht-kultivierter hoch-affiner MB (Bender und Conrad, 1992; King, 1993; Roslev et al., 1997) ist allerdings nicht unumstritten. Dunfield gelang der Nachweis, daß der $K_{\text{M(app)}}$ von *Methylobacter* sp. Stamm LR1 nach mindestens 10-tägiger Inkubation bei niedrigen Methan-Konzentrationen von zunächst 12,6 μM auf bis zu 56 nM abfiel (Dunfield and Conrad, 2000). Wenngleich *Methylobacter* sp. Stamm LR1 unter diesen Bedingungen kein Wachstum sondern lediglich einen Erhaltungsstoffwechsel zeigte, bietet diese Studie eine Alternative zum Konzept der obligat hoch-affinen MB.

Es sind somit drei Konzepte zur Erklärung der hoch-affinen Methan-Oxidation zu unterscheiden:

- 1) Die Existenz bislang nicht isolierter oder charakterisierter MB mit einer starken Anpassung an niedrige Methan-Konzentrationen. Die in Böden mit hohen Methan-Aufnahmeraten häufig detektierten *pmoA*-Sequenz von USC α und USC γ wurden als ein deutlicher Hinweis auf solche hoch-affinen MB gewertet (Holmes et al., 1999; Knief et al., 2003b).
- 2) Die Fähigkeit mancher Methanotropher, unter Methan-Mangel von niedriger auf hohe Affinität umstellen zu können. Das einige MB dieses Potential haben, wurde am Beispiel *Methylocystis* sp. Stamm LR1 gezeigt. Es ist allerdings fraglich, ob die von Stamm LR1 gezeigte Affinität ausreicht, um die in Böden gemessenen, zum Teil deutlich geringeren $K_{M(app)}$ -Werte zu erklären. Darüberhinaus wurde bislang kein *Wachstum* unter atmosphärischen Methankonzentrationen nachgewiesen.
- 3) Eine Synthese der beiden geschilderten Konzepte, nämlich die Koexistenz von obligat und fakultativ hoch-affinen MB in einem Habitat.

Das Vorhandensein und auch die hohe Abundanz von ungewöhnlichen *pmoA*-Sequenzen in als Methan-Senken beschriebenen Böden läßt alleine noch keinen hinreichenden Schluß auf den physiologischen Zustand der von diesen Sequenzen repräsentierten Gruppe zu. Zwecks Identifikation der physiologisch aktiven MB führten deshalb verschiedene Autoren Markierungsversuche mit ^{13}C - bzw. ^{14}C -markiertem Methan durch (Bull et al., 2000; Holmes et al., 1999; Knief et al., 2003; Roslev and Iversen, 1999). Während der Inkubation wurde das markierte Substrat von den MB oxidiert und der Kohlenstoff teilweise assimiliert. Die Analyse der auf diese Weise markierten PFLAs erlaubt Rückschlüsse auf die zugehörigen physiologisch aktiven Organismen, da methanotrophe Gattungen anhand ihrer charakteristischen Hauptfettsäuren identifiziert und differenziert werden können (Bowman et al., 1991). Bei den in verschiedenen Inkubations-Studien markierten Fettsäuren handelt es sich hauptsächlich um C18-Fettsäuren. Ebenfalls markierte C17-Fettsäuren sind in bekannten MB nur in Spuren vorhanden, was von den Autoren als ein weiterer Hinweis auf nicht-kultivierte, aktive MB gewertet wurde. C18-Fettsäuren sind charakteristisch für methanotrophe *Alphaproteobacteria* (Bull et al., 2000; Holmes et al., 1999a; Roslev und Iversen, 1999). Da die bei phylogenetischen Berechnungen innerhalb der *Alphaproteobacteria* gruppierenden *pmoA*-Sequenzen des USC α in den

gleichen bzw. vergleichbaren Böden mit hoher Abundanz nachgewiesenen werden konnten, wurde vermutet, daß die markierten Fettsäuren zu dieser Gruppe gehören könnten, und daß diese MB den aktiven Teil der methanotrophen Gemeinschaft darstellen (Bull et al., 2000; Roslev und Iversen 1999). Eine vergleichbare Analyse wurde von Knief und Mitarbeitern (2003) für Böden durchgeführt, in denen überwiegend Sequenzen des USC γ nachgewiesen wurden. Tatsächlich gelang bei allen untersuchten Böden der Nachweis von für *Gammaproteobacteria* charakteristischen PLFAs. Auch hier folgerten die Autoren einen unmittelbaren Zusammenhang zwischen den durch die USC γ -Sequenzen charakterisierten Organismen und der für diese Böden nachgewiesenen hoch-affinen Methan-Oxidation.

1.3 Der Kohlenstoff-Stoffwechsel und die pMMO methanotropher Bakterien

Der initiale Schritt der Umsetzung von Methan zu CO₂ ist die Oxidation des Methans zu Methanol bei gleichzeitiger Reduktion molekularen Sauerstoffs zu H₂O. Die Reaktion wird durch die lösliche Methan-Monooxygenase (sMMO) oder durch die membrangebundene Form, die pMMO katalysiert. Die sMMO wurde bei Vertretern aller Gattungen methanotropher *Alphaproteobacteria* gefunden, darüber hinaus kommt sie bei Vertretern der Gattungen *Methylomonas*, *Methylococcus* und *Methylobacterium* vor (Dedysh et al., 2000; Fuse et al., 1998; Heyer et al., 2002a; Stainthorpe et al., 1990). Dennoch kann man von der phylogenetischen Stellung eines MB nicht auf das Vorhandensein der sMMO schließen. Insbesondere viele Typ I MB, aber auch manche Typ II MB haben keine sMMO. Im Gegensatz dazu ist die pMMO bei den bekannten MB mit Ausnahme der Gattung *Methylocella* universell vorhanden. Die benötigten Reduktionsäquivalente werden der sMMO durch Oxidation von NADH + H⁺ zur Verfügung gestellt, während das Reduktionsmittel der pMMO noch nicht eindeutig geklärt werden konnte (vgl. 1.6; Chan et al., 2004; Zahn und DiSpirito, 1996, Choi et al., 2003). Das Enzym Methanol-Dehydrogenase oxidiert das im ersten Schritt entstandene Methanol zu Formaldehyd. Die bei der Methanol-Oxidation frei werdenden Reduktionsäquivalente werden mittels des Kofaktors

Pyrroloquinolin-Chinon der Atmungskette zugeführt (Goodwin und Anthony, 1998). Die klassische Annahme ist, daß die Oxidation von Formaldehyd zu Formiat durch Formaldehyd-Dehydrogenasen unter Spaltung von H_2O katalysiert werden kann, wobei die Elektronen auf NAD(P)^+ übertragen werden (Hanson und Hanson, 1996). Im Jahre 2001 wurde eine neuartige, membranassoziierte Form der Formaldehyd-Dehydrogenase in *M. capsulatus* (Bath) nachgewiesen, die den oder den Cytochrom-bc1-Komplex als Elektronen-Akzeptor nutzt (Zahn et al., 2001). Bei der durch Formiat-Dehydrogenase katalysierten Oxidation des Formiats freiwerdende Reduktionsäquivalente werden direkt auf NAD^+ übertragen.

Alternativ zur Oxidation können MB das Formaldehyd (alle MB) oder auch Formiat (Typ X MB) als Kohlenstoffquelle nutzen und assimilationistisch verstoffwechseln. Typ I MB nutzen dazu den Ribulose-Monophosphat-Weg (RuMP-Weg), wobei Ribulose-Monophosphat als C1-Akzeptor mit Formaldehyd zu Hexulose-6-Phosphat kondensiert. Bei drei Durchläufen des zyklischen Weges werden so drei Mol Formaldehyd zu einem C3-Körper umgesetzt. Beim Serin-Weg der Typ II MB reagiert Formaldehyd mit Glycin zu Serin. Insgesamt werden zwei Mol Formaldehyd und ein Mol CO_2 zu einem C3-Körper verbunden. Für Typ X MB konnte zusätzlich der Benson-Calvin-Zyklus der CO_2 -Fixierung nachgewiesen werden. CO_2 wird an Ribulose-Diphosphat gebunden, das instabile Produkt zerfällt in zwei Moleküle Phosphoglycerinsäure. Insgesamt verschmelzen 3 Moleküle CO_2 zu einem C3-Körper (Hanson und Hanson, 1996).

Neben dem vorgestellten Weg der Formaldehyd-Oxidation wurden in den letzten Jahren für eine Vielzahl methylotropher Bakterien zusätzliche Stoffwechselwege beschrieben. Chistoserdova und Mitarbeiter (2000) zeigten, daß weitere dissimilatorische Varianten der Formaldehyd-Oxidation existieren. Es wurden verschiedene lineare Formaldehyd-Stoffwechselwege identifiziert, die sich hauptsächlich durch den jeweils verwendeten Kofaktor unterscheiden. Als wichtigste C1-Carrier sind zu nennen: Tetrahydrofolat (H_4F), Tetrahydromethanopterin (H_4MPT) und Glutathion (Vorholt et al; 2002). In Vertretern (*Methylosinus* spp., *Methylococcus* spp., *Methylomonas* spp.) aller drei MB-Typen wurde der H_4MPT -abhängige Oxidationsweg nachgewiesen. Zu Enzymen dieses Weges hochähnliche homologe Enzyme wurden ursprünglich für methanogene und sulfatreduzierende *Archaea* nachgewiesen. Es wird vermutet, daß sie vermittelt horizontalen

Gentransfers von den *Proteobacteria* erworben wurden (Vorholt et al; 2002). Die genaue Rolle der einzelnen Wege im Kohlenstoff-Stoffwechsel ist Gegenstand aktueller Forschung. Neben der dissimilatorischen Rolle (Oxidation von Formaldehyd zu Formiat) und einer assimilierenden Funktion (H_4F dient als C1-Träger im Serin-Weg) ist vor allem die Detoxifikation intermediärer Produkte Gegenstand der Diskussion.

1.4 Methan-Monooxygenasen

Verfügt ein Organismus über beide Formen der MMO wird die Expression durch die Kupfer-Konzentration des Habitats bestimmt. Konzentrationen über ca. 1 nmol Cu pro mg Zellprotein induzieren die Expression der pMMO, während bei niedrigeren Konzentrationen die sMMO exprimiert wird (Murrell et al., 2000, Choi et al., 2004). Die alleinige Expression der sMMO resultiert in geringeren Wachstumserträgen, was durch das im Vergleich zur pMMO (vgl. 1.6) möglicherweise negativere Redoxpotential des Elektronen-Donators der sMMO ($\text{NADH} + \text{H}^+$) und eine dadurch bedingte ungünstigere Energiebilanz erklärt wurde (Hanson and Hanson, 1996).

pMMO und sMMO sind evolutionär unabhängig voneinander entstanden, sie sind nicht homolog. Die sMMO besteht aus einer Reduktase (MMOR), einer regulatorischen Komponente "B" (MMB), und einer Hydroxylase-Komponente (MMOH). MMOH enthält zwei dinukleare Eisen-Zentren und ist als Hexamer in $\alpha_2\beta_2\gamma_2$ -Konfiguration organisiert. Die sMMO wird von fünf Genen eines Gen-Clusters kodiert: MMOH wird von *mmoX*, *mmoY*, und *mmoZ*, MMOR und MMB durch *mmoC* and *mmoB* kodiert (Shigematsu et al., 1999). Im Vergleich zur pMMO verfügt die sMMO über ein außerordentlich breites Substratspektrum. Wichtigstes, neben dem eigentlichen Wachstumssubstrat kooxidiertes Edukt ist Trichloroethylen (TCE), andere Substrate sind beispielsweise Chlor-Methan, Toluol (Methylbenzol), Chloroform oder Mono- und Dichlor-Ethan. In Abwesenheit von Methan werden diese Umweltgifte als einzige Quelle von Kohlenstoff und Energie genutzt (Lipscomb, 1994). Das Substratspektrum der pMMO ist wesentlich eingeschränkter. Lediglich kurzkettige Alkane werden hydroxyliert, kurzkettige Alkene epoxidiert. Darüberhinaus sind die Kooxidationsraten wesentlich niedriger. Die TCE-

Oxidationsraten der pMMO liegen beispielsweise um zwei bis drei Größenordnungen unter denen der sMMO (Lontoh et al., 2002; Oldenhuis et al., 1992).

Die pMMO ist homolog zu der ebenfalls membranassoziierten Ammonium-Monooxygenase (AMO) der Ammoniak-oxidierenden Bakterien (Holmes et al., 1995). Sie besteht aus einer Hydroxylase und einer Oxidoreduktase. Dem in der molekularen Mikrobiologie weit verbreiteten Sprachgebrauch folgend, wird innerhalb dieser Dissertation mit „pMMO“ die Hydroxylase-Untereinheit bezeichnet. Die Hydroxylase besteht aus 3 Polypeptiden: PmoC (γ -Untereinheit), PmoA (β -Untereinheit), und PmoB (α -Untereinheit) in der Stöchiometrie $(\alpha\beta\gamma)_2$. (Zahn und DiSpirito, 1996; Choi et al., 2003). Sie werden von den Genen *pmoC*, *pmoA* und *pmoB* kodiert, die in den bislang daraufhin untersuchten MB (2 Typ II MB, Gilbert et al., 2000; 1 Typ X MB, Stolyar et al., 1999) gemeinsam in einem Operon reguliert werden. Stolyar und Mitarbeiter (1999) sowie Klotz und Norton konnten zeigen, daß häufig zwei hochähnliche Kopien des pMMO- bzw. AMO-Operons in einem Genom vorhanden sind (Klotz and Norton, 1998). So unterscheiden sich die beiden Kopien von *M. capsulatus* Bath nur an 13 von 3183 Positionen (Stolyar et al., 1999). "Knock-out"-Experimente deuten darauf hin, daß die Produkte beider Operon-Kopien funktionell redundant sind und einander substituieren können (Stolyar et al., 2001a). In den Genomen einiger MB bzw. Ammoniak-Oxidierer befindet sich zusätzlich eine weitere *pmoC*-/*amoC*-Kopie mit bislang unbekannter Funktion.

Wenngleich bereits seit den siebziger Jahren des vergangenen Jahrhunderts versucht wird, die Struktur der pMMO durch biochemische und biophysikalische Methoden zu bestimmen, scheiterten die Bemühungen bislang daran, daß es nicht gelingt, das Enzym mittels Detergentien aus der Membran zu lösen, ohne es dabei zu denaturieren (vgl. 1.5). Die spezifische Bindung des "Suizid"-Substrats Acetylen an die PmoA-Untereinheit läßt darauf schließen, daß diese das aktive Zentrum trägt. Entsprechende Befunde liegen auch für das Homologon AmoA vor (Zahn and DiSpirito, 1996).

1.5 Biochemische Charakterisierung der partikulären Methan-Monooxygenase

Die Reaktionen beider Methan-Monooxygenasen, der partikulären wie auch der löslichen Form, sind prinzipiell ähnlich und entsprechen der Reaktion klassischer Monooxygenasen (Goirl und Dooley, 1999). Zwei Reduktionsäquivalente werden genutzt, um die O-O-Bindung molekularen Sauerstoffs zu spalten. Ein Sauerstoffatom wird dabei zu H₂O reduziert. Das zweite Sauerstoffatom wird in die C-H-Bindung des Methans eingefügt, wobei Methanol entsteht. Während die sMMO biochemisch gut charakterisiert werden konnte, scheiterten Versuche, die funktionelle und räumliche Struktur der pMMO aufzuklären. 1996 gelang es Nguyen und Mitarbeitern erstmals, die pMMO so weit aufzureinigen, daß eine Auftrennung aller drei Untereinheiten mittels SDS-Page und eine anschließende Sequenzierung der N-terminalen Enden mittels Edmann-Abbau möglich wurde (Nguyen et al., 1996). Bereits 1995 erfolgte nach N-terminaler Sequenzierung der PmoB-Untereinheit die erstmalige Identifizierung der *pmo*-Gene (Semrau et al., 1995a). Molekulare Massen, die *in silico* für von den Gensequenzen abgeleitete Polypeptide berechnet wurden, stehen in Übereinstimmung mit durch MALDI-TOF-Massenspektrometrie bestimmten molekularen Massen der drei Polypeptide (42, 29 und 28 kDa).

Bisher war es nicht möglich, die pMMO aus dem IZM-System zu lösen, ohne dabei die Enzymstruktur zumindest partiell zu zerstören. Die erzielten Enzympräparationen wiesen lediglich geringe, zum Teil keine Aktivität mehr auf, was auf a) Denaturierung des Enzyms während der Präparation, b) Verlust des Reduktionsmittels und/oder c) Verlust des Aktivators zurückgeführt wurde (Chan et al., 2004; Tikhvatullin et al., 2001). So sind oft stark von einander abweichende, zum Teil widersprüchliche Ergebnisse publiziert worden (Basu et al., 2003; Nguyen et al., 1998; Takeguchi et al., 1998; Zahn und DiSpirito, 1996). Zwei bis 21 Kupferatome und null bis drei "Nicht-Häm"-Eisenatome wurden für verschiedene pMMO-Präparationen berichtet. Das Vorhandensein von Cu²⁺- und Fe²⁺-Ionen wurde sowohl durch Mössbauer-Spektroskopie wie auch mittels Atom-Absorptions-Spektroskopie bestätigt (Chan et al., 2004; Choi et al., 2003; Lieberman et al., 2004; Semrau et al., 1995b; Zahn and DiSpirito, 1996). Vergleichbare Probleme wurden auch für die homologe AMO Ammoniak-oxidierender Bakterien berichtet (Zahn et al., 1996). Auch das für die Reduktion molekularen Sauerstoffs von der pMMO benötigte

Reduktionsmittel konnte nicht zweifelsfrei bestimmt werden. Zwei konkurrierende Modelle (benötigte Reduktionsmittel Quinon bzw. $\text{NADH} + \text{H}^+$) werden diskutiert (Basu et al., 2003; Chou et al., 2003; Ngugen et al., 1998). Kürzlich gelang Yu und Mitarbeitern die Aufreinigung eines homogenen pMMO-Präparates von *M. capsulatus* Bath. Sie konnten 13-14 Kupfer-Atome pro 99 kDa Monomer ($\alpha\beta\gamma$) nachweisen (Yu et al., 2003)..

1.6 Eine neuartige paraloge Genkopie der *pmoA*

Die für die pMMO kodierenden Gene *pmoC*, *pmoA* und *pmoB* sind in den Genomen der bislang daraufhin untersuchten MB als Operon organisiert. Verschiedene MB weisen eine zweite, nahezu Sequenz-identische Kopie des *pmo*-Operons auf (vgl. 1.5). Es war deshalb überraschend, daß Dunfield und Mitarbeiter (2002) in dem neu isolierten Typ II MB *Methylocystis* sp. Stamm SC2 zwei verschiedene, stark von einander abweichende *pmoA*-Fragmente nachweisen konnten (prozentuale Ähnlichkeit auf Gen-Ebene 73%, auf Aminosäure-Ebene 68,5%). Die Reinheit der Kultur wurde durch mikroskopische Untersuchungen, mittels „Fluoreszenz-In-Situ-Hybridisierung“ (FISH) sowie durch Sequenzierung der *mxoF*- und 16S rRNA-Gene nachgewiesen, sodaß beide *pmoA*-Fragmente zweifelsfrei einem Organismus zugeordnet werden konnten. Eines der beiden Fragmente gruppierte bei phylogenetischen Verrechnungen gemeinsam mit *pmoA*-Sequenzen anderer Stämme der Gattung *Methylocystis* und reflektierte somit die phylogenetische Position von *Methylocystis* sp. Stamm SC2. Die zweite Sequenz gruppierte jedoch gemeinsam mit einer zuvor aus Reisboden isolierten Umweltsequenz (Horz et al., 2001), getrennt von den *pmoA*- und *amoA*-Sequenzen bekannter Organismen.

Um einen ersten Eindruck von der Verbreitung der neuartigen Genkopie innerhalb der Gruppe methanoxidierender Bakterien zu gewinnen, wurde ein spezifischer PCR-Assay entwickelt und auf DNA-Extrakte von zehn Typ II MB angewendet. Dadurch konnte die Existenz einer zweiten, abweichenden *pmoA*-Genkopie auch für einen Vertreter der Gattung *Methylosinus* (*M. trichosporium* Stamm KS21) nachgewiesen werden. Die zusätzlichen Gen-Kopien, im Folgenden "novel" *pmoA* oder auch *pmoA2* genannt, gruppierten gemeinsam, unabhängig von

der phylogenetischen Stellung der zugehörigen Organismen und formten eine eigenständige Linie innerhalb der aus *Alphaproteobacteria* isolierten *pmoA*-Sequenzen. Ob *pmoA2* Teil eines kompletten Operons ist, und insbesondere ob dieses Operon für ein funktionell aktives Enzym kodiert oder ob es sich bei *pmoA2* möglicherweise lediglich um nicht exprimierte, funktionslose DNA-Artefakte handelt, wurde in dieser Studie nicht untersucht (Dunfield et al., 2002). Auch eine detailliertere Charakterisierung der Verbreitung der *pmoA2*, insbesondere bei Typ I und Typ X MB, erfolgte zunächst nicht.

1.7 Der Stickstoff-Stoffwechsel methanotropher Bakterien

Die Fähigkeit, molekularen Stickstoff als einzige Stickstoffquelle nutzen zu können, ist eine wichtige physiologische Eigenschaft vieler MB. So zeigten Vertreter der Gattungen *Methylococcus* (Typ X), *Methylocystis*, *Methylosinus* (Typ II), *Methylocella* und *Methylocapsa* Wachstum auf N-freiem Medium (Murrell und Dalton, 1983; Dedysh et al., 2000; Auman et al., 2001; Dedysh et al., 2002). Lediglich für Typ I MB konnte bislang kein Wachstum auf N-freien Medien nachgewiesen werden (vgl. 3.3 und 4.3), wenngleich die notwendigen Gene im Genom vorhanden sind. Allerdings wird diese Befähigung zur N₂-Fixierung bei den meisten Gattungen durch erhöhte Konzentrationen molekularen Sauerstoffs limitiert. Unter hohem O₂-Partialdruck zeigen Vertreter der Typ II und Typ X MB sowie der Gattung *Methylocella* nur geringes Wachstum auf N-freien Medien. Abnehmender O₂-Partialdruck resultiert sowohl in erhöhtem Wachstum, wie auch in erhöhter Acetylen-Reduktions-Aktivität dieser Organismen (Auman et al., 2001; Dedysh et al., 2000 ; Murrell und Dalton, 1983). Die Messung der Acetylen-Reduktion dient dem Nachweis der Aktivität der Dinitrogenase (Auman et al., 2001). *M. acidiphila* hingegen ist zu exponentiellem Wachstum auf N-freien Medien sowohl unter aeroben wie auch mikroaeroben Bedingungen befähigt (Dedysh et al., 2002) und erinnert darin stark an die auf Basis der 16S-rRNA-Gene nahe verwandten heterotrophen Organismen der Gattung *Beijerinckia*.

Die biogene N₂-Fixierung wird durch das Enzymsystem Nitrogenase katalysiert. Die Nitrogenase methanotropher Bakterien ist ein oligomerer Komplex

und besteht aus zwei Komponenten: Dinitrogenase (auch Nitrogenase-Eisen-Protein genannt) katalysiert die Reduktion von Stickstoff zu Ammonium und wird durch das Gen *nifD* kodiert. Die zweite Komponente, Dinitrogenase-Reduktase, ist ein Homodimer des Proteins NifH und überträgt mittels zweier 4Fe-4S-Kluster Elektronen von Ferredoxin auf das Eisen-Protein (Pau, 1989; Georgiadis et al., 1992). Eine Analyse der *nifH*-Gene von vier Typ I und sieben Typ II MB ergab hohe Ähnlichkeiten (94-99%) zu vielen zuvor publizierten Umwelt-Sequenzen, was als ein wichtiges Indiz für die weite Verbreitung N₂-Fixierender MB und ihren Einfluß auf den N-Kreislauf gewertet wurde (Auman et al., 2001). Ren und Mitarbeiter (2000) konnten darüber hinaus anhand eines *Methylobacter*-Stammes nachweisen, das manche MB unter aeroben Bedingungen Stickstoffmonoxid (NO) produzieren, unter anaeroben Bedingungen jedoch verbrauchen, wobei Distickstoffoxid (N₂O) entsteht.

1.8 Ziele der Arbeit

Der Fokus dieser Arbeit lag auf methanotrophen Bakterien mit ungewöhnlichen *pmo*-Genen. Die in dieser Arbeit dargestellten Versuche sind jedoch zwei von einander abgegrenzten Themengebieten zuzuordnen. Zum einen sollten Fragen hinsichtlich der neuartigen paralogen Kopie der *pmoA* (*pmoA2*) methanotropher Bakterien beantwortet werden. Andererseits waren die bisher nicht kultivierbaren, potentiell hoch-affinen MB des USCα Gegenstand der Arbeit. Der methodische Schwerpunkt lag zum einen in der Etablierung der Methoden der sogenannten Metagenomik. Andererseits sollten Kenntnisse der Bioinformatik erarbeitet und hinsichtlich der verschiedenen Fragestellungen angewendet werden.

Mittels spezifischer PCR-Assays sollte zunächst die Verbreitung der *pmoA2* innerhalb der MB untersucht und anschließend durch vergleichende phylogenetische Analysen erste Erkenntnisse über die Evolution dieser Gen-Kopie gewonnen werden. Anschließend sollte anhand vergleichender Untersuchungen der Operon-Strukturen von pMMO-1 und pMMO-2 des Referenz-Organismus *Methylocystis* sp. Stamm SC2 die grundlegende Frage beantwortet werden, ob neben *pmoA2* auch ungewöhnliche Gen-Kopien von *pmoB* und *pmoC* existieren. Im positiven Fall sollte die Arbeit

Aussagen hinsichtlich Expression und Funktion der pMMO-2 ermöglichen und so die Grundlage für Expressionsstudien sowie "Knockout"-Experimente legen.

Die bei der Erstellung einer Genom-Bibliothek von *Methylocystis* sp. Stamm SC2 zu erwerbende Methoden-Kompetenz sollte anschließend für die Arbeiten über USC α genutzt werden. Mit Hilfe des sogenannten Metagenom-Ansatzes sollten größere Genom-Fragmente aus einem als Methan-Senke charakterisierten Waldboden erzielt und erste Einblicke in das Genom dieser bislang nicht kultivierbaren MB gewonnen werden. Im Gegensatz zur üblichen Vorgehensweise der Umwelt-Genomik sollten hierbei gezielt Sequenz-Informationen einer nicht angereicherten Organismengruppe gewonnen werden. Ziel dieser Arbeiten war zum einen die Aufklärung und Analyse der Operonstruktur der für pMMO kodierenden Gene, sowie die vergleichende Interpretation der daraus abgeleiteten Polypeptide hinsichtlich ungewöhnlicher Eigenschaften, welche die mögliche hohe Affinität des USC α zu Methan erklären könnten. Darüberhinaus sollte die Frage der phylogenetischen Zuordnung des USC α durch bioinformatische Auswertung der gewonnenen Sequenzdaten befriedigend beantwortet werden. Ein weiteres, in der geschilderten Analyse des USC α enthaltenes Ziel dieser Arbeit war die Erstellung einer großen Metagenom-Bibliothek aus Waldboden-DNA, die auch für andere Fragestellungen genutzt werden könnte. Zu Referenzzwecken sollte eine genomische Bibliothek von *M. acidiphila* erstellt werden und der das *pmoA*-Operon enthaltender Genomabschnitt sequenziert und analysiert werden.

Eine zunächst ebenfalls beabsichtigte Sequenzierung größerer genomischer Abschnitte eines USC α -Vertreters mittels der gezielten Suche nach überlappenden Fragmenten mußte aufgrund der geringen Abundanz des USC α aufgegeben werden. Die dadurch freigesetzten Ressourcen wurden für andere Arbeiten über methanotrophe Bakterien, die Entwicklung bioinformatischer Werkzeuge sowie Metagenomik-basierte Studien unkultivierter Vertreter der Domäne *Archaea* (Daten in dieser Arbeit nicht gezeigt) genutzt.

2. Material und Methoden

2.1 Chemikalien

Soweit nicht anders vermerkt, wurden sämtliche Chemikalien in p.A. - Qualität von den Firmen Boehringer Mannheim (Mannheim), Fluka (Buchs, Schweiz), Merck (Darmstadt) und Sigma-Aldrich (Seelze) bezogen. Enzyme sowie andere molekularbiologische Reagenzien wurden von Amersham (Braunschweig), Applied Biosystems (Weiterstadt), Biozym (Oldendorf), New England Biolabs (Frankfurt) und Promega (Mannheim) bezogen.

2.2 Sterilisationsverfahren

Alle im Rahmen der vorliegenden Arbeit verwendeten Reaktionsgefäße, Pipettenspitzen, Lösungen, Puffer und sonstige Materialien wurden, soweit nicht durch den Hersteller sterilisiert, durch Autoklavieren (121°C, 20 min) sterilisiert. Lediglich die Robotter-Platten (vgl. 2.6.9) wurden durch UV-Bestrahlung sterilisiert.

2.3 Verwendete Organismen und Referenzstämme

Im Rahmen dieser Arbeit wurden die in Tabelle 2.1 aufgeführten Organismen verwendet.

Tabelle 2.1: Verwendete Mikroorganismen

Stamm	Herkunft
<i>Methylocystis</i> sp. SM16	Heyer et al. (2002)
<i>Methylocystis</i> sp. LR1	Heyer et al. (2002)
<i>Methylocystis</i> sp. B2/7	Heyer et al. (2002)
<i>Methylocystis</i> sp. 62/12	Heyer et al. (2002)
<i>Methylocystis</i> sp. KS9	Heyer et al. (2002)
<i>Methylocystis</i> sp. 014c	Heyer et al. (2002)
<i>Methylocystis</i> sp. IMET 10484	Heyer et al. (2002)
<i>Methylocystis</i> sp. 21/1	Heyer et al. (2002)
<i>Methylocystis</i> sp. IMET 10499	Heyer et al. (2002)
<i>Methylocystis</i> sp. F10v2a	Heyer et al. (2002)
<i>Methylocystis</i> sp. KS12	Heyer et al. (2002)

Material und Methoden

<i>Methylocystis</i> sp. IMET 10486	Heyer et al. (2002)
<i>Methylocystis</i> sp. SC2	Heyer et al. (2002)
<i>Methylocystis</i> sp. Pi5/4	Heyer et al. (2002)
<i>Methylocystis parvus</i> 81	P. Dunfield
<i>Methylosinus sporium</i> SC8	Heyer et al. (2002)
<i>Methylosinus sporium</i> SK13	Heyer et al. (2002)
<i>Methylosinus sporium</i> 20/3	Heyer et al. (2002)
<i>Methylosinus sporium</i> H1b	Heyer et al. (2002)
<i>Methylosinus trichosporium</i> SM6	Heyer et al. (2002)
<i>Methylosinus trichosporium</i> H3	Heyer et al. (2002)
<i>Methylosinus trichosporium</i> M23	Heyer et al. (2002)
<i>Methylosinus trichosporium</i> 39/3	Heyer et al. (2002)
<i>Methylosinus trichosporium</i> OB3b	NCIMB 11131 ¹
<i>Methylosinus trichosporium</i> SC10	Heyer et al. (2002)
<i>Methylosinus trichosporium</i> KS21	Heyer et al. (2002)
<i>Methylomonas</i> sp. D1a	P. Dunfield
<i>Methylocaldum</i> sp. E10a	P. Dunfield
<i>Methylococcus capsulatus</i> Bath	NCIMB 11853 ¹
<i>Methylocapsa acidiphila</i> B2	DSM 13967 ¹ , Dedysh et al.(2004)
<i>Methylocella palustris</i> K	ATCC 700799 ¹ , Dedysh et al.(2004)
<i>Methylocella palustris</i> Ch3	Dedysh et al.(2004)
<i>Methylocella palustris</i> Y5	Dedysh et al.(2004)
<i>Methylocella palustris</i> H4	Dedysh et al.(2004)
<i>Methylocella silvestris</i> BL2	DSM 15510 ¹ , Dedysh et al.(2004)
<i>Methylocella tundrae</i> T4	Dedysh et al.(2004)
<i>Methylocella tundrae</i> Y1	Dedysh et al.(2004)
<i>Methylocella tundrae</i> Ch1	Dedysh et al.(2004)
<i>Methylococcus capsulatus</i> Texas	NCIMB 11853 ¹
<i>Methylobacter luteus</i>	NCIMB 11914 ¹
<i>Methylomicrobium album</i>	NCIMB 11123 ¹
<i>Methylosinus trichosporium</i> OB3b	ATCC 35070 ¹
<i>Methylosinus sporium</i> SM27a	Heyer et al. (2002)
<i>Methylosinus sporium</i> 44/2	Heyer et al. (2002)
<i>Methylocystis echinoides</i>	Heyer et al. (2002)
<i>Methylocystis</i> sp. H9a	Heyer et al. (2002)
<i>Beijerinckia indica</i> subsp. <i>indica</i>	ATCC 9039 ¹
<i>Beijerinckia indica</i> subsp. <i>lacticogenes</i>	DSM 1719 ¹
<i>Beijerinckia mobilis</i>	DSM 2326 ¹
<i>Beijerinckia derxii</i> subsp. <i>derxii</i>	DSM 2328 ¹
<i>Beijerinckia derxii</i> subsp. <i>venezuelae</i>	DSM 2329 ¹
<i>Rhodoblastus acidophila</i>	ATCC 25092 ¹
<i>Azorhizobium caulinodans</i>	DSM 5975 ¹
<i>Bradyrhizobium japonicum</i>	DSM 30131 ¹
<i>E.coli</i> ElektroMax® DH10BTM	Invitrogen Life Technologies (Carlsbad, USA)
<i>E.coli</i> EPI300™ Plating strain	Epicentre (Fosmid Library Production Kit)
<i>E.coli</i> LE392MP Control strain	Epicentre (Fosmid Library Production Kit)
<i>E.coli</i> Klon mit <i>pmoA</i> USCα	S. Kolb (Zit)

Abkürzungen: ATCC: American Type Culture Collection, Rockville, USA; NCIMB: National Collections of Industrial and Marine Bacteria Ltd, Aberdeen, GB; DSM: Deutsche Sammlung für Mikroorganismen und Zellkulturen GmbH, Braunschweig

2.4 Anzucht von Mikroorganismen

2.4.1 Anzucht methanotropher Bakterien

Methanotrophe Bakterien wurden unter Standardbedingungen auf „Medium 10“ (Heyer et al., 2002), pH 7,3 bei 30°C auf dem Horizontalschüttler (150 upm) unter einer Gasphase mit 17% CH₄, 21% O₂ und 0,3% CO₂ angezogen. Die Begasung erfolgte mittels einer Gasmischanlage oder durch manuelle Injektion des benötigten Volumens CH₄. Das Gas wurde mittels eines 0,2 µm Filters (Schleicher und Schuell, Dassel) sterilfiltriert. Angeimpft wurde jeweils 5%ig mit Stamm- bzw. Vorkulturen. Das Wachstum wurde durch Messung von Zelldichte (OD₆₀₀) und Methan-Konzentration verfolgt. Bei Bedarf wurde die Gasphase während des Wachstums ausgetauscht bzw. CH₄ oder O₂ nachgegeben.

2.4.2 Anzucht von Klonen

Sämtliche Klonierungsarbeiten wurden mit *E.coli*-Zellen durchgeführt. Die Inkubationstemperatur betrug für Platten und Flüssigmedium 37°C, als Medium wurde Luria-Bertani-Medium (LB, 1% [w/v] Trypton, 1% [w/v] NaCl, 0,5% [w/v] Hefeextrakt) verwendet. Je nach Antibiotika-Resistenzgen des verwendeten Vektors wurden dem Medium Chloramphenicol (cam, 6,25 µg ml⁻¹) oder Kanamycin (kan, 50 µg ml⁻¹) zugesetzt. Alle Arbeiten wurden unter oxischen Bedingungen durchgeführt.

2.5 Verwendeter Boden

Für die Erstellung der Metagenom-Bibliothek aus Waldboden-DNA wurde ein hinsichtlich der Population methanotropher Bakterien gut untersuchtes (Kolb et al., 2003; Knief et al., 2003b; Henckel et al., 2000) Buchenwaldgebiet (*Fagus*, *Quercus*) nahe Marburg beprobt. Der Bodentyp dieses Systems ist Braunerde auf Buntsandstein, die Textur wurde als Sandiger Sand klassifiziert (Kolb, 2004). Im Mittel wurde ein pH-Wert von pH = 4,3 bestimmt.

Verrottende Blätter sowie die aufliegende Humusschicht wurden entfernt, und die Probe dem oberen mineralischen Horizont (A_h , 1-15 cm Tiefe) entnommen. In diesem Bereich wurden die höchsten Methan-Oxidationsraten gemessen (Knief, 2004). Die Proben wurden in 50 ml Polyvinylchlorid-Gefäßen ("Falcon-Tubes", Kobe, Marburg) bei -20°C gelagert und maximal zweimal aufgetaut. Insgesamt wurden zwei Beprobungen durchgeführt (20.10.2003 und 19.04.2004).

2.6 Erstellung von "Large Insert Libraries"

2.6.1 "Large Insert Libraries"

Herkömmliche Verfahren der DNA-Extraktion nutzen häufig physikalische Methoden des Zellaufschluß („Bead Beating“, Ultraschall, „French Press“ u.a.). Aufgrund dabei auftretender Scheerkräfte liegt die Größe der isolierten DNA durchschnittlich im Bereich von 5 bis max. 20 Kilobasen (kb) (Zhou et al., 1999). Für die Erstellung von Large Insert Libraries (LILs) sind die so gewonnenen Fragmente zu klein. Deshalb wurden in dieser Arbeit für die Erstellung von BAC- und Fosmid-Bibliotheken ausschließlich „sanfte“, chemische Lysemethoden verwendet. Ein gravierender Nachteil solcher Methoden kann, je nach bakterieller Zielgruppe, darin liegen, daß während der sanften Lyse nur ein Teil der in der Probe vorhandenen Mikroorganismen aufgeschlossen wird, was in einer kontraproduktiven physikalischen Selektion (Abreicherung der Zielgruppe) resultieren kann. Während der Optimierung der Methoden wurde deshalb neben der erzielbaren Fragmentgröße auch die relative Aufschlußrate potentieller Zielgruppen in einem semiquantitativen PCR-gestützten Ansatz verfolgt (Daten nicht gezeigt).

Bakterielle Artifiziale Chromosome (BACs) sind für die Klonierung großer DNA-Fragmente ("Inserts") hervorragend geeignet. BACs können Inserts mit Größen von vielen hundert kb aufnehmen. In dieser Arbeit kamen zwei BAC-Vektoren zum Einsatz, pBeloBAC11 (Epicentre Technologies, Biozym, Hessisch Oldendorf) und pIndigoBac5 ("Cloning-Ready"; Epicentre). Der 7,4 kb große Vektor pBeloBAC11 (Abb. 2.1) gehört der zweiten Generation von BAC-Vektoren an und wurde aus pBAC108L (Shizuya *et al.*, 1992) abgeleitet (Kim *et al.*, 1996). Die Zielsequenzen der für das Linearisieren des Vektors zu verwendenden Restriktionsenzyme

("Multiple Cloning Site") befinden sich innerhalb eines *lacZ*-Gens, daß von T7 und SP6 Promotoren flankiert wird. Dadurch wird die Überprüfung der Effizienz einer erfolgten Ligation mittels Blau-Weiß-Test ermöglicht. Ein Chloramphenicol-Resistenzgen erlaubt die gezielte Selektion transformierter Zellen. Als weiterer BAC-Vektor wurde pIndigoBAC-5 (Abb. 2.1, Birren et al., 1999) verwendet. Der Vektor wird „ready-for-cloning“ angeboten, ist daher bereits durch Restriktion mit *Hind*III oder *Bam*HI linearisiert und dephosphoryliert. Alternativ zu der Klonierung mittels BACs (vgl. 2.6.2 und 2.6.6) wurde der Fosmid-Vektor pCC1FOS™ in Kombination mit "MaxPlax™ Lambda Packaging Extracts" (beides Epicentre) für die Erstellung von LILs eingesetzt (Abb 2.1; vgl. 2.6.3 und 2.6.7).

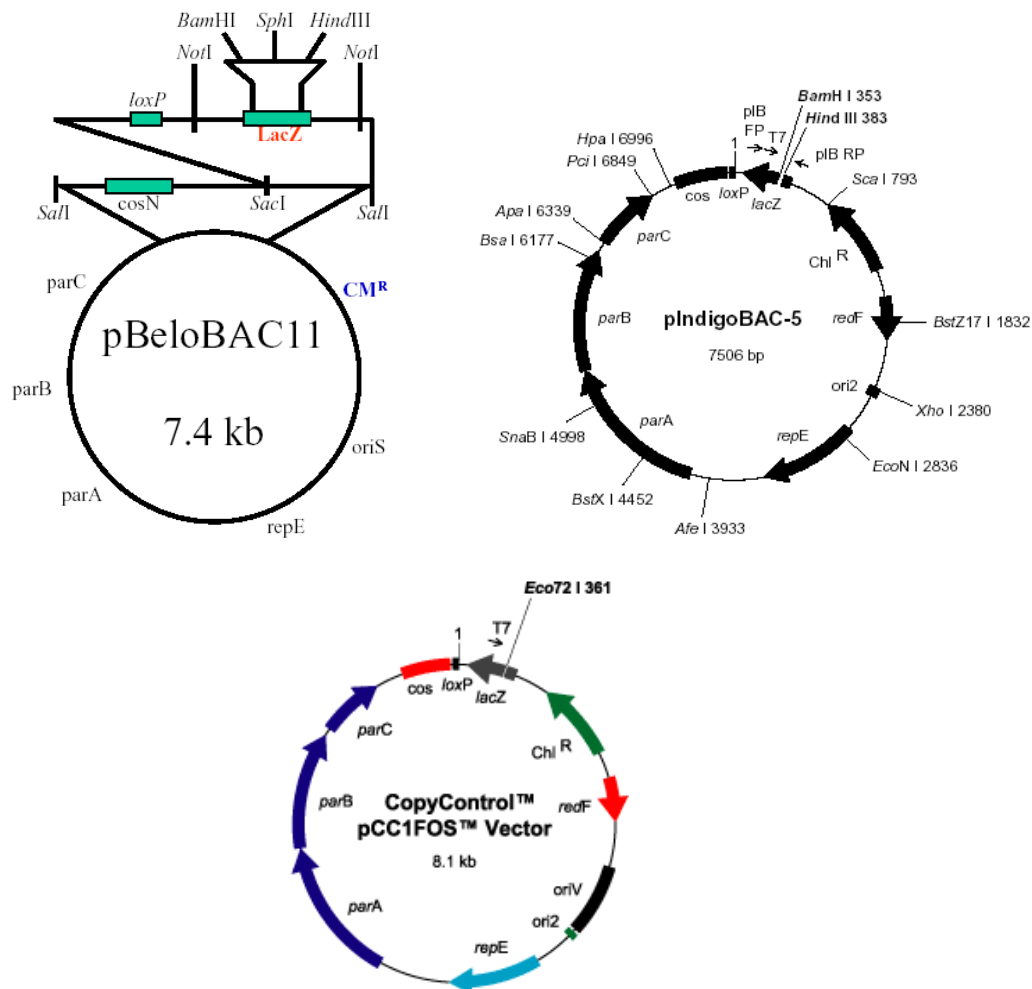


Abb. 2.1 Karten der in dieser Arbeit verwendeten HMW-Vektoren. BAC-Vektoren : pBeloBAC11 und pIndigoBAC-5; Fosmid-Vektor : pCC1FOS™ Vector CM = Chl = Chloramphenicol-Resistenzgen. Nach: Kim et al., 1996 (pBeloBAC), Birren et al., 1999 (pIndigoBAC) sowie Fa. Epicentre (pCC1FOS™).

2.6.2 DNA-Präparation für die Erstellung von BAC-Banken

Einbettung in Agarose

Für die Extraktion hochmolekularer (high molecular weight, HMW) DNA mit Fragmentgrößen von mehr als 100 kb ist es notwendig, auf die DNA einwirkende Scheerkräfte auf ein Minimum zu reduzieren. Dazu gehören auch relativ geringe Kräfte, wie sie beim Pipettieren von Flüssigkeiten auftreten. Beim Ansatz der „Einbettung in Agarose“ wird deshalb versucht, die DNA möglichst nicht in Flüssigkeiten zu prozessieren. Stattdessen wird sie während der aufeinanderfolgenden Schritte Zellaufschluß, DNA-Aufreinigung, partielle Restriktion und Größenselektion in Agarose eingebettet (Peterson et al., 2000; Ochsenreiter et al., 2002).

Die wie unter 2.4 beschrieben angezogenen Zellen wurden bei 7000 g pelletiert, der Überstand verworfen und die Zellen in Puffer (10 mM Tris, 20 mM NaCl, 50 mM EDTA, pH 8,0; Konzentration: 1×10^{10} Zellen ml^{-1} Puffer) resuspendiert. Die Zellen wurden auf 45°C erwärmt und mit Aliquots 2x konzentrierter, auf 45°C abgekühlter flüssiger Low Melting Point (LMP) Agarose (PE) vermischt. Das Agarose-Zellen-Gemisch wurde in Aliquots von 100 μl in Formen verfüllt, wo es erkalten und sich verfestigen konnte. Die Endkonzentration der eingebetteten Zellen betrug somit 5×10^9 Zellen ml^{-1} 1x konzentrierter Agarose.

Zellaufschluß und Aufreinigung der DNA

2 ml Agarose-Zellen-Gemisch (20 Agarosewürfel) wurden für 20 h bei 37° C in Lysis-Puffer (100 mM Tris-HCl, 100 mM EDTA, 1,5 M NaCl, 1% [w/v] Hexadecyl-Trimethylammonium-Bromid, 2% [w/v] Natrium-Dodecylsulfat; pH 8,0; Dedysh et al., 1998a) inkubiert, mit 25 ml Waschpuffer (WP, 20 mM Tris-HCl, 50 mM EDTA; pH 8,0) gewaschen und anschließend in 5 ml Proteinase K-Puffer (100 mM EDTA, 0,2% Natrium-Deoxycholat, 1% Natrium-Lauryl-Sarcosin, 1 mg ml^{-1} Proteinase K; pH 8,0) für weitere 14 h bei 50°C inkubiert. Überreste abgebauter Proteine wurden durch viermaliges Waschen in WP für jeweils 1 h bei Raumtemperatur entfernt. Anschließend wurden die DNA-haltigen Agarosewürfel in 0,5 M EDTA (pH 8,0) bei 4° C bis zur Durchführung nachfolgender Schritte gelagert.

Partielle Restriktion und Größenselektion

Für die Klonierung in BACs ist es erforderlich, mittels Restriktionsenzymen an Vektor und zu klonierender DNA überhängende Enden zu erzeugen. Für pBeloBAC11 wurde die Endonuklease *Hind*III, für pIndigoBac5 das Enzym *Bam*HI genutzt. Während pIndigoBAC bereits seitens des Herstellers linearisiert ist, mußte für Arbeiten mit pBeloBAC11 nicht nur die genomische DNA sondern auch der Vektor selbst mit Endonuklease behandelt werden.

Zur Produktion größerer Mengen von pBeloBAC11 wurde der Vektor mit einem herkömmlichen pUC-Plasmid über eine *Bam*HI-Schnittstelle fusioniert, was die Produktion als „High-Copy“-Vektor in *E. coli* erlaubte (Daten nicht gezeigt). Nach Extraktion (QIAGEN Plasmid Maxi Kit, QIAGEN, Hilden) und weiterer Aufreinigung mit Plasmid-SafeTM ATP-Dependend DNase (Epicentre) wurde das Konstrukt mit *Bam*HI geschnitten und BAC-Vektor und pUC-Plasmid durch präparative Agarose-Gelelektrophorese getrennt. Nach Extraktion der BAC-DNA aus dem Agarosegel (QIAquick Gel Extraction Kit, QIAGEN, Hilden) wurde der Vektor dephosphoryliert (HK Phosphatase, Epicentre) und im Anschluss auf mögliche Selbstligation überprüft.

Vor der Nuklease-Behandlung (partiell schneiden) der genomischen DNA sind Vorstudien für jeden neuen DNA-Extrakt notwendig. Ziel dieser Studien ist die Bestimmung einer geeigneten Menge Restriktionsenzym, die in definierter Inkubationszeit die genomische DNA so partiell schneidet, daß ein maximaler Teil der DNA im gewünschten Größenbereich liegt. Da die Enzymmenge von der durchschnittlichen Größe und Konzentration der jeweiligen DNA abhängt, mußte die Voruntersuchung für jedes DNA-Extrakt erneut durchgeführt werden. 10 bis 20 DNA-haltige Agarosewürfel (insgesamt 1-2 ml) wurden zunächst für 1 h bei 50° C in 50 ml Puffer A (0,1% [w/v] Ascorbinsäure, 2,0% [w/v] PVP-40, and 0,13% [w/v] Natriumdiethyldithiocarbamat, 0,5 M EDTA, pH 9.1), anschließend für 1 h in 50 ml Puffer B (0,1% [w/v] Ascorbinsäure, 2,0% [w/v] PVP-40, 13% [w/v] Natriumdiethyldithiocarbamat, 0,05 M EDTA, pH 8.0) auf Eis und daraufhin für 30 min in 50 ml Puffer C (0,1% [w/v] Ascorbinsäure, 2,0% [w/v] PVP-40, and 0,13% [w/v] Natriumdiethyldithiocarbamat, 10 mM EDTA, pH 8.0, 10 mM Tris-HCl, pH 8.0) auf Eis inkubiert. Im Anschluß wurde der Puffer gegen frischen Puffer C ausgetauscht und 50 µl 0,1 M Phenylmethansulfonyl-Fluorid (PMSF) zugegeben, um von der DNA-Extraktion möglicherweise zurückgebliebene Spuren von Proteinase K

zu inhibieren. Nach Inkubation für 1 h auf Eis wurde der letzte Schritt wiederholt und der Ansatz für eine weitere Stunde auf Eis inkubiert. PMSF sowie Reste des in hohen Konzentrationen in Puffer A vorhandenen EDTAs wurden durch mehrmaliges Austauschen des Puffers C bei fortgesetzter Inkubation auf Eis entfernt. 3 100µl-Agarosewürfel wurden anschließend für 1 h in Puffer D (10% 10x *Hind*III Puffer, 1,0 mM DTT, 4 mM Spermidin, 1,0% [v/v] 100x BSA) auf Eis äquilibriert, geviertelt und in 1,5 ml Eppendorfcups mit jeweils 250 µl Puffer D überführt. Die im Text verwandte Bezeichnung *Hind*III ist sinngemäß durch *Bam*HI zu ersetzen. *Hind*III (20 U µl⁻¹) wurde mit 9 Teilen Wasser (Sigma-Aldrich) und einem Teil 10x *Hind*III-Puffer verdünnt und anschließend in definierten Mengen zu den Agarosewürfeln gegeben. Eine einstündige Inkubation auf Eis ermöglichte das Einwandern des Enzyms in die Agarosewürfel bei minimaler Enzymaktivität. Anschließend wurden die Reaktionsansätze für genau 20 min bei 37°C inkubiert, erneut auf Eis überführt und die Reaktion mit 30 µl 0,5 M EDTA (pH 8,0) gestoppt. Auftrennung der partiell geschnittenen DNA auf dem Pulsfeldgel (Abb. 2.2) erlaubte die Bestimmung der angemessenen Enzymkonzentration. Anschließend wurden die verbleibenden Würfel auf die selbe Art mit der so bestimmten optimalen Enzymmenge behandelt.

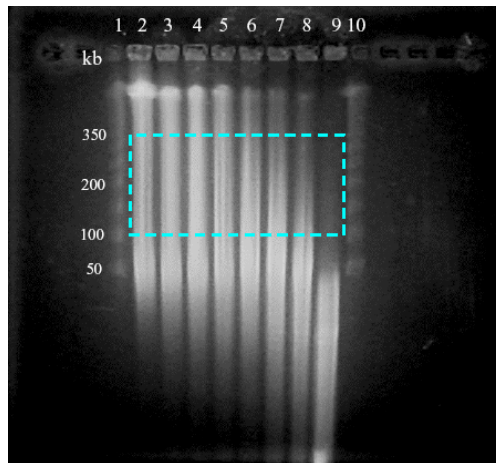


Abb. 2.2 : Partielle Restriktion genomischer DNA, Vorstudie. Aliquots genomischer DNA werden mit verschiedenen Konzentrationen einer Endonuklease für definierte Zeit inkubiert. Zu geringe Enzymmengen resultieren in einem hohen Anteil ungeschnittener DNA (Spuren 2,3), zu hohe Konzentrationen führen zu vollständiger Restriktion (Spur 9). Die optimale Enzymkonzentration wird bestimmt (Spur 4) und anschließend für die partielle Restriktion der restlichen DNA ("Large Scale") eingesetzt. Größenmarker : Mid-Range PFG-Marker (NEB).

2.6.3 DNA-Präparation für die Erstellung von Fosmid-Banken

DNA-Extraktion aus Waldboden

Jeweils 4 g Waldboden wurden mit 3 ml 0,12 M Natrium-Phosphat-Puffer (pH 8,0) und 250 µl Lysozym-Stocklösung (100 mg ml⁻¹) verrührt und bei 37°C inkubiert. Nach 1 h erfolgte die Zugabe von jeweils 60 µl Proteinase K (20 mg ml⁻¹). Nach einer weiteren Stunde wurden die Proben in flüssigem Stickstoff gefroren und anschließend im Wasserbad bei 65° C aufgetaut. Sie verblieben nach Zugabe von jeweils 400 µl 20% SDS-Lösung für eine Stunde bei 65° C. Die Proben wurden zunächst 5 min bei 3.500 g abzentrifugiert, die Überstände anschließend bei 10.000 g für 1 min erneut zentrifugiert. Die Präzipitate wurden verworfen, die Überstände weiter aufgereinigt. Um die auftretenden Scheerkräfte zu minimieren, wurden alle, auch die während der nachfolgend beschriebenen Schritte notwendigen Pipettierarbeiten mit abgeschnittenen Pipettenspitzen durchgeführt.

Aufreinigung von Umweltproben

Zur Aufreinigung hochmolekularer DNA aus Umweltextrakten wurde das FastDNA[®]SPIN[®] Kit for Soil (Q-BIOgene) genutzt. Die wie oben beschrieben extrahierte DNA wurde zunächst mit 0,45 Volumenteilen 6 M Kalium Acetat vermischt und für 5 min bei Raumtemperatur inkubiert. Proteinreste wurden durch Zentrifugation für 1 min bei 10.000 g in einer Mikrozentrifuge präzipitiert. Die Überstände wurden, wie vom Hersteller angegeben, auf die im FastDNA[®] SPIN[®] Kit (Q-BIOgene) enthaltenen Säulen geladen. Dazu wurden jeweils 3 ml klaren Überstandes mit 2 ml "Binding-Matrix" in ein 15 ml Polypropylenröhrchen (Kobe, Marburg) gegeben und die DNA durch dreiminütiges vorsichtiges Schwenken an die Silikamatrix gebunden. Die Matrix wurde im Anschluss auf zwei Säulenröhrchen (Spin Colum, Q-BIOgene) gegeben, der enthaltene Flüssigkeitsanteil durch Zentrifugation für 1 min bei 10.000 g in einer Mikrozentrifuge abgeschieden. Zusätzlich zu dem vom Hersteller empfohlenen Aufreinigungsschritt wurden die Säulen 3 bis 4 mal mit jeweils 500 µl 5,5 M Guanidin Thiocyanat gewaschen. Die weitere Aufreinigung erfolgte nach den Angaben der Fa. Q-BIOgene durch Waschen mit SEWS-M (Ethanol-haltige Salzlösung), Trocknen der Matrix und Elution der DNA von jeder Säule mit 100 µl Wasser (Sigma-Aldrich). Alle Wasch- und Elutionsschritte wurden, soweit nicht anders angegeben, durch Zentrifugation in einer

Mikrozentrifuge bei 13.000 upm für 1 min durchgeführt. Die DNA wurde im Anschluß durch Vakuum-Zentrifugation für 30 – 45 min 3-fach ankonzentriert. Der Erfolg der DNA-Extraktion wurde durch Auftrennung auf einem präparativen Pulsfeldgel kontrolliert.

Enzymatische Reparatur der Enden gescheerter DNA

Bei Verwendung von Fosmiden für das Erstellen von LILs ist, im Gegensatz zur Verwendung der BAC-Technologie, keine Behandlung der genomischen DNA zur Erzeugung überhängender Enden notwendig. Da somit die bei der partiellen Restriktion auftretende Verringerung der durchschnittlichen Fragmentgröße (partiell schneiden von 300-400 kb großen Fragmenten resultiert in Fragmenten der Größenordnung 50 bis 150 kb) umgangen wird, kann auch während der Extraktion stark fragmentierte DNA (40 kb) erfolgreich in Fosmide kloniert werden. Vor der Größenselektion sind jedoch die durch Einwirkung von Scheerkräften entstandenen überhängenden einzelsträngigen Enden der DNA aufzufüllen.

52 µl DNA-Extrakt wurden auf Eis platziert, anschließend 8 µl 10x „End-Repair“-Puffer, 8 µl 2,5 mM dNTP-Mix, 8 µl 10 mM ATP und 4 µl „End-Repair Enzyme Mix“ (alle Substanzen: Copy ControlTM Fosmid Library Produktion Kit, Epicentre) zugegeben und gut verrührt. Nach 1 h Inkubation bei RT wurde das Enzymgemisch 10 min lang bei 70° C hitzeinaktiviert.

2.6.4 Größenselektion mittels PFGE

Hochmolekulare DNA wurde mittels der Pulsfeld-Gelelektrophorese (PFGE) aufgetrennt. Wichtigstes Charakteristikum der PFGE im Vergleich zur konventionellen eindimensionalen Gelelektrophorese ist die periodisch wechselnde Laufrichtung der DNA. Die Ausrichtung der am Pulsfeldgel anliegenden Spannung wechselt dabei mit frei wählbarer Frequenz. Die Frequenzlänge bestimmt dabei das Auftrennungspotential der PFGE-Laufs. Große DNA-Moleküle brauchen längere Zeit, um sich in Richtung der neuen Spannungsquelle zu orientieren. Bei hinlänglich hoher Frequenz wird große DNA nahezu auf der Stelle bewegt. Kürzere Moleküle orientieren sich schneller neu und laufen im Gel dementsprechend rascher.

Während dieser Arbeit wurden ausschließlich 1%ige Pulsfeldgele (1,5 g Pulse Field Agarose (Bio-Rad) in 150 ml 0,5x TBE Puffer (0,05 M Tris, 1 mM EDTA, pH 8,0, 0,05 M Borsäure) verwendet.. Der Winkel zwischen den beiden Laufrichtungen betrug bei allen während dieser Arbeit durchgeführten PFGE-Läufen 120°. Die Dauer eines Laufs sowie die Frequenzen des Spannungsquellen-Wechsels wurden an die aufzutrennende DNA angepaßt.

Bei der Erstellung von BAC-Bibliotheken werden sowohl während der Ligation als auch während der anschließenden Transformation kleinere DNA-Moleküle bevorzugt. Schon geringe Mengen kurzer DNA-Fragmente verhindern deshalb die effiziente Klonierung hochmolekularer DNA. Da die hinreichende Abtrennung niedermolekularer DNA durch einfache Größenselektion mittels PFGE nicht gewährleistet ist, wurde zur Konstruktion von BAC-Bibliotheken eine Dreischritte-Auftrennung durchgeführt (modifiziert nach Oseogava et al., 1998). Während des ersten Schritts wird eine umgekehrte Spannung an das Gel angelegt (3 h, Frequenz : 15s, - 6 V cm⁻¹). Die DNA wandert entgegen der sonst üblichen Laufrichtung. Kleine Moleküle laufen dabei aus dem Gel in den umgebenden Puffer aus. Während des zweiten Schritts migriert die hochmolekulare, im Gel verbliebene DNA in die Ausgangsposition zurück (3 h, Frequenz : 15s, + 6 V cm⁻¹). Gemeinsam mit einem im Anschluß an Schritt 2 aufgetragenen Größenstandard (MidRange PFG Marker I, NEB) wird sie im dritten Schritt aufgetrennt (18 h, Frequenzen : 1-40s, +6 V cm⁻¹). Da bei Einsatz der Fosmid-Technik aufgrund der verwendeten Phagen ohnehin nur DNA-Fragmente im Bereich von 35 bis 45 kb kloniert werden können, konnten hier die Schritte 1 und 2 entfallen. Die Auftrennung der DNA erfolgte in einem Lauf (15 h, Frequenzen : 1-12s, 6 V cm⁻¹).

Um nach erfolgter Klonierung die durchschnittliche Insertgröße einer LIL zu ermitteln, wurden ungeschnittene, wie unter 2.6.8 angegeben extrahierte, hochmolekulare Plasmide für 10 h bei Frequenzen von 1-10 s und einer Spannung von 6 V cm⁻¹ aufgetrennt.

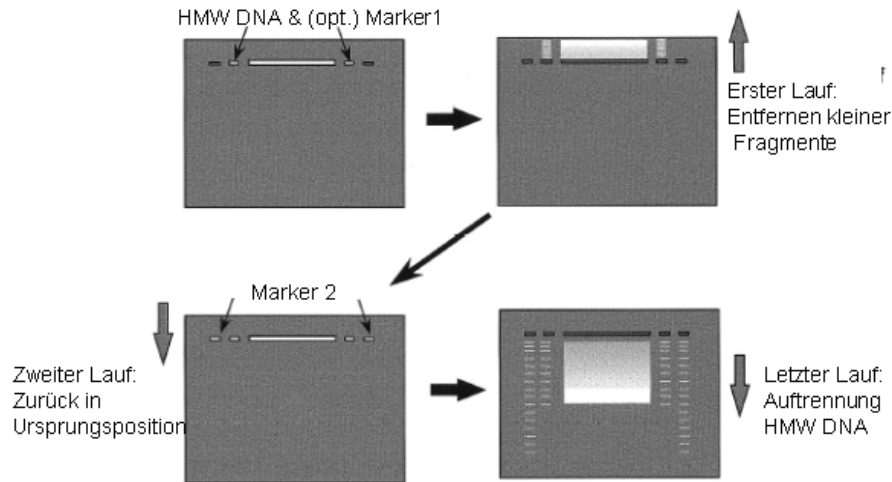


Abb. 2.3 : Größenselektion partiell geschnittener genomischer DNA. Präferenziell klonierte kleine DNA-Fragmente (< 25 kb) wurden durch PFGE mit umgekehrter Laufrichtung vor der eigentlichen Größenselektion entfernt. Abbildung modifiziert nach Oseogawa und Mitarbeiter (1998).

2.6.5 Elution größenselektierter DNA aus dem Agarosegel

Verschiedene Methoden der Elution von DNA aus Agarosegelen sind in der Literatur beschrieben. Während dieser Arbeit wurde neben der im Folgenden dargestellten Elution mittels Agarase auch die Elektroelution angewendet, jedoch mit deutlich geringerer Ausbeute. Andere Methoden (z.B. Verwendung der Kits der Fa. QIAGEN) resultierten in sehr starker Degradierung der DNA durch Scheerung. Es wurden keine für die Arbeit relevanten Ergebnisse mittels dieser Methoden erzielt. Die präparative Aufbereitung von in Agarose gebundenen Nukleinsäuren mittels enzymatischen Abbaus der Agarose ist nur möglich, wenn die Agarose einen niedrigen Schmelzpunkt (Low melting point agarose, LMPA) aufweist. Nur diese Agarose kann bei den für die verwendeten Enzyme optimalen Temperaturen in flüssigem Zustand gehalten werden, was für einen effizienten Abbau unerlässlich ist. War die anschließende Extraktion der größenselektierten DNA aus dem Gel geplant, wurde deshalb der zentrale Bereich des Pulsfeldgels durch 1% LMPA (in 0,5% TBE) ersetzt. Nach der erfolgten Größenselektion wurden nur die Randbereiche der aufgetrennten DNA sowie die flankierenden Marker in Ethidiumbromid (EtBr) - Lösung angefärbt. EtBr interkaliert in doppelsträngige DNA und verändert dabei sein Anregungs-Spektrum. Durch Anregung mit UV-Licht kann so die im Gel gebundene DNA sichtbar gemacht

werden. Anhand der verwendeten Größenstandards wurde der Bereich des Gels, der DNA gewünschter Größe enthielt, ausgewählt und markiert. Dadurch war es möglich, den zentralen Bereich des Gels, der den Großteil aufgetrennter genomischer DNA gewünschter Größe enthielt, zu markieren, ohne diese DNA EtBr bzw. UV-Strahlung auszusetzen. Der ausgewählte Bereich wurde mit einem Skalpell ausgeschnitten und mit einem Gemisch aus β -Agarase (NEB) und Gelase (Epicentre) wie folgt behandelt. Die Agarose wurde mit 1/10 Vol 10x β -Agarase Puffer (NEB) versetzt und bei 65° C geschmolzen. Anschließend wurde sie auf eine Temperatur von 42° C abgekühlt. Bei dieser Temperatur bleibt LMPA flüssig. Pro 100 mg LMPA wurde 1 U β -Agarase sowie 1 U „Gelase Enzymmix“ zugegeben und vorsichtig vermischt. Nach Inkubation bei 42° C für 1 h wurde das Enzym-Gemisch bei 70° C für 10 min hitzeinaktiviert. Während 5 min Inkubation auf Eis wurden unvollständig abgebaute Agarosereste verfestigt und anschließend für 20 min bei 13.000 upm in einer Mikrozentrifuge präzipitiert. Der die DNA enthaltende Überstand wurde abgenommen, mittels Vakuum-Zentrifugation ankonzentriert und bis zum Gebrauch bei 4° C gelagert.

2.6.6 Klonierung mittels BAC-Vektoren

Ligation

Geschnittene, größenselektierte und aufgereingte DNA wurde über Nacht mit dem jeweils verwendeten BAC-Vektor ligiert. Da die DNA-Konzentrationen nach Aufreinigung stets deutlich unter der von der Firma Epicentre empfohlenen Konzentration lag, wurde die DNA mittels Vakuumzentrifugation ankonzentriert. Die Bestimmung der DNA-Konzentration erfolgte photometrisch (E_{260}) bzw. auf dem Agarosegel. Stets wurden Ansätze mit verschiedenen Verhältnissen [w/w] von Vektor- und Insert-DNA parallel pipettiert (berechnete Verhältnisse: 1:1; 1:5; 1:10; 1:20). Genomische DNA, Vektor und 10x Ligase Puffer (NEB) wurden auf Eis gelagert, bis Vektor und Puffer aufgetaut waren. Die Reaktionen wurden in Gesamtvolumina von 50 bis 150 μ l angesetzt. Zu 1/10 Gesamtvolumen 10x Puffer wurden ca. 100 - 150 ng genomische DNA, eine dem gewünschten Verhältnis entsprechende Menge Vektor sowie 1-3 μ l T4-Ligase (400 U μ l⁻¹, NEB) gegeben, durch vorsichtiges Fingerschnippen vermengt und über Nacht bei 16° C inkubiert.

Anschließend wurde das Enzym für 30 min bei 65° C hitzeinaktiviert und das Ligationsgemisch wie unten beschrieben prozessiert.

Transformation

Vor der Elektroporation wurde das Ligationsgemisch mittels Mikrodialyse entsalzen. Dazu wurde ein 0,025 µm Nitrozellulosefilter (Millipore, Schwalbach) vorsichtig auf 30 ml sterilem 20% [v/v] Polyetylen glykol appliziert. Das Reaktionsgemisch wurde mit abgeschnittener Pipettenspitze in der Mitte des Filters aufgetragen und das Dialysegefäß verschlossen. Nach 3h Inkubation bei Raumtemperatur wurde das entsalzene Reaktionsgemisch vorsichtig abgenommen und bis zur Transformation bei 4° C gelagert. Jeweils 2 bis 4 µl Ligationsansatz wurden mit 20 µl kompetenter Zellen (*E. coli* ElektroMax[®] DH10B[™], Invitrogen) verrührt, auf Eis gekühlt und in ebenfalls gekühlten 1 mm Elektroporationsküvetten (PeqLab Biotechnology GmbH, Erlangen) bei 2,5 kV, 25 µF und 100 Ω in einem Elektroporator (Bio-Rad *E. coli* Pulser, Bio-Rad Laboratories, Hercules, USA) transformiert.

2.6.7 Klonierung mittels Fosmiden

Für die Konstruktion von LILs aus Umwelt-DNA wurde die Fosmid-Technologie sowie das CopyControl[™] Fosmid Library Production Kit (Epicentre) genutzt. Nach enzymatischer Rekonstruktion der doppelsträngigen DNA-Enden ("End-Repair") gröbenselektierte (durchschnittlich 40 kb), mittels Vakuumzentrifugation ankonzentrierte DNA wurde gemäß folgendem Protokoll mit dem Vektor pCC1FOS[™] ligiert. 1 µl 10x "Fast-Link" Ligations-Puffer, 1 µl 10 mM ATP, 1 µl pCC1FOS[™] Vektor (0,5 µg/µl), 6 µl DNA und 1 µl Fast-Link-DNA-Ligase (Alle Substanzen Copy Control[™] Fosmid Library Produktion Kit, Epicentre) wurden auf Eis vermischt und 2 bis 3 h bei RT inkubiert. In manchen Fällen erfolgte eine zusätzliche Inkubation über Nacht bei 4° C. Das Enzym wurde im Anschluß für 10 min bei 70° C hitzeinaktiviert.

Die Klonierung erfolgte nach einem modifizierten Protokoll der Fa. Epicentre. Ein Aliquot (50 µl) MaxPlax Lambda Packaging Extracts (Copy Control[™] Fosmid Library Produktion Kit, Epicentre) wurde auf Eis aufgetaut, 25 µl entnommen, mit 10 µl Ligationsansatz vermischt und der Rest unmittelbar bei -80° C gefroren. Nach Inkubation für 2 h bei RT wurden die restlichen 25 µl auf Eis getaut und zu dem

ersten Ansatz gegeben. Nach weiteren 2 h bei RT wurden 940 µl „Phage Dilution“-Puffer (10 mM Tris/HCl, pH 8,3; 100 mM NaCl; 10 mM MgCl₂) sowie 25 µl Chloroform zugegeben. Die in Phagen verpackte DNA konnte einige Tage bei 4°C gelagert werden. Bei längerer Lagerung nahm die Transformations-Effizienz stark ab.

Für die Transformation wurden jeweils 100µl verdünnter Phagenpartikel mit 1 ml *E.coli* EPI300TM Kultur (OD₆₀₀ = 1) vermischt, für 20 min bei 37°C inkubiert und anschließend in Aliquots von 200 µl auf mit LBcam-Agarplatten ausplattiert. Nach Inkubation über Nacht konnte die Transformationseffizienz durch Auszählen der Kolonien abgeschätzt werden.

2.6.8 Größenanalyse der BAC- und Fosmid-Inserts von Klonen

Um den Erfolg der Klonierung zu überprüfen, insbesondere um die gemittelte Größe der klonierten DNA-Fragmente festzustellen, wurden 10 bis 20 Klone jeder neu erstellten Klonbibliothek zufällig ausgewählt, die Plasmide (BACs oder Fosmide) extrahiert und die Insertgröße mittels PFGE bestimmt.

Plasmid-Extraktion : Methode 1 [Modifiziert nach Peterson et al. (2000)]

4 ml der über Nacht gewachsenen Kultur (vgl. 2.4) wurden bei 3000 g für 5 min pelletiert, der klare Überstand verworfen, die pelletierten Zellen in 200 µl Puffer (50 mM Glukose, 25 mM TrisHCl, pH 8,0, 10 mM EDTA) resuspendiert, in ein 1,5 ml Eppendorfcup überführt und für 5 min auf Eis inkubiert. Nach Zugabe von 400 µl Lysis-Puffer (0,2 M NaOH, 1% [w/v] SDS) und Durchmischen durch mehrmaliges Invertieren der Gefäße erfolgte eine weitere Inkubation auf Eis für 5 min. Nach Zugabe des Neutralisations-Puffers (3,0 M Kalium-Acetat; 28,5% [v/v] Essigsäure) und Inkubation für 10 min auf Eis wurden die ausgefallenen Proteine und Zellreste für 25 min bei 10.000 g in einer Mikrozentrifuge pelletiert. Der DNA-haltige Überstand wurde anschließend vorsichtig abgenommen und mit 600 µl eiskalten Isopropanols vermischt. Nach Inkubation für 20 min bei -80° C wurde die DNA 30 min bei 10.000 g ausgefällt, der Überstand verworfen und das Pellet zweimal mit eiskaltem 70% EtOH gewaschen. Die Plasmid-DNA wurde 1-2 h unter der Clean-Bench getrocknet und über Nacht bei 4° C in 50 µl Wasser (Sigma-Aldrich) gelöst.

Plasmid-Extraktion : Methode 2

Alternativ zu der beschriebenen zeitaufwendigen aber kostengünstigen Methode wurde das QIAprep[®] Spin Miniprep Kit (QIAGEN) nach Herstellerangaben (Protokoll für Mikrozentrifuge) verwendet. Für die Präparation großer Mengen DNA wurden die Kits „Qiaprep[®] Plasmid Midi Kit“ und „Qiaprep[®] Plasmid Maxi Kit“ (beide QIAGEN) nach Herstellerangaben benutzt.

Größenbestimmung mittels *NotI* oder *XhoI*-Restriktion

Die Auftrennung ungeschnittener, zu einem großen Teil zirkulär vorliegender hochmolekularer Plasmide im Pulsefeldgel läßt häufig nur eine ungefähre Abschätzung der Insertgrößen zu. Für eine genauere Bestimmung wurden jeweils ca. 2,5 µg der extrahierten BACs bzw. Fosmide mit einem Restriktionsenzym geschnitten. Dazu wurden 1 µl Enzym-spezifischen 10x Puffers, 1 µl BSA (1 µg µl⁻¹) und 15 U Restriktionsenzym mit 5 µl extrahierter DNA und 2 µl Wasser (Sigma-Aldrich) verrührt und für 2 h bei 37° C inkubiert. Anschließend wurde das Enzym durch 30- minütige Inkubation bei 65° C hitzeinaktiviert. Die geschnittene DNA wurde auf einem 0,6%igen Agarosegel 3 h aufgetrennt. Als Größenstandard dienten Smart Ladder (Eurogentec, Seraing, Belgien) und mit *HindIII* geschnittene DNA des λ-Phagen (NEB) (vgl. 2.6.4).

2.6.9 Vervielfältigung, Screening und Lagerung von LILs

Im Verlauf dieser Arbeit wurden zwei hinsichtlich Zielsetzung und Effizienz stark von einander abweichende "Screening"-Techniken angewendet. Für eine dauerhafte, gesicherte Aufbewahrung der Klonbibliotheken wurden Kolonien "gepickt" und die Klone in Mikrotiterplatten angezogen. Die Mikrotiterplatten wurden repliziert, und die Replikate langfristig bei -70° C eingelagert. Für ein einmaliges Hochdurchsatz-Screening wurde das zeitaufwändige „Picken“ umgangen. Stattdessen wurden die Agarplatten mittels einer Stempel-Technik direkt kopiert. Nach vollzogenem Screening wurden alle Platten verworfen, da eine dauerhafte Lagerung auf Agarplatten nicht möglich ist.

Lagerung und Screening von Mikrotiterplatten

Die relativ kleinen, aus genomischer DNA von *M. acidiphila* bzw. *Methylocystis* sp. Stamm SC2 erstellten BAC-Bibliotheken (jeweils ca. 1000 Klone) wurden zur dauerhaften Aufbewahrung manuell von den Agarplatte in 10 bzw. 11 Mikrotiterplatten (96 Kulturen in je 100 µl LB cam) überführt ("Picken"). Nach Inkubation über Nacht wurden die Flüssigkulturen mittels eines Mikrotiterplatten-Replikators (Eigenbau) in „Freeze-Medium“ (LBcam mit 4,4% [v/v] Glycerol, 1,7 mM MgCl, 6,8 mM NH₄2SO₄, 13 mM KH₂PO₄, 36 mM K₂HPO₄, 0,4 mM MgSO₄) überführt. Die Originalplatten wurden für das PCR-basierte Screening verwendet, die Replikate bei -70° C langfristig eingelagert. Für das Screening wurden jeweils 12 in einer Mikrotiterplatten-Reihe wachsende Klone zu sogen. „Reihen-Poolen“ kombiniert. Aliquots der 8 Reihen-Poole einer Platte wiederum wurden zu „Platten-Poolen“ vereinigt. Zur Identifikation von Klonen, deren Inserts für die pMMO bzw. andere Zielgene kodieren, wurden zunächst Platten-Poole mit entsprechenden spezifischen PCR-Assays (vgl. 2.7.6) untersucht, anschließend die Reihen-Poole der Test-positiven Platten. Um die gesuchte Test-positive Einzelkolonie zu identifizieren wurden abschließend die jeweils 12 Klone einer Test-positiven Reihen gescreent. Positive Kolonien wurden mehrfach durch Ausstreichen vereinzelt und auf Reinheit getestet, bevor sie Verwendung in weiteren Untersuchungen fanden. Die Zellen von Poolen wie Einzelkolonien wurden durch zehnminütiges Kochen bei 100° C lysiert, die Zellreste anschließend 5 min bei 7000 g pelletiert und 1 µl des klaren Überstandes direkt in die PCR-Reaktion eingesetzt.

Agarplattenscreening

Die aus rund 250.000 Klonen zusammengesetzte Metagenom-Bibliothek wurde lediglich partiell in Mikrotiterplatten überführt. Dazu konnte ein Roboter des MPI für molekulare Genetik, Berlin genutzt werden. Ausplattieren auf großen, an die Robotik angepassten Agarplatten (20 x 20 cm) resultierte in durchschnittlich 2.000 Kolonien Platte⁻¹. Die Erkennungs-Effizienz des Roboters lag bei ca. 90%. Insgesamt wurden 30.000 Klone langfristig eingelagert. Der überwiegende Teil der Fosmid-Banken wurden nicht archiviert. Statt dessen wurden die auf Agarplatten gewachsenen Kolonien mittels eines mit Samt überzogenen Stempels (Eigenkonstruktion) auf andere Agarplatten übertragen (repliziert). Benutzter Samt wurde autoklaviert, gewaschen, getrocknet und vor erneutem Gebrauch durch Autoklavieren sterilisiert.

Jeweils eine Kopie jeder Agarplatte wurde bei 4° C für einige Wochen gelagert, die andere Kopie diente dem Screening. Dazu wurden die 200-400 Kolonien pro Platte mit sterilem Wasser abgeschwämmt, und als „Agarplatten-Poole“ gelagert. 200 µl jeden Pools wurden mit einem Aliquot Glycerol-Puffer (65% [v/v] Glycerol, 1,2 mM MgSO₄ pH 8,0) vermischt und in flüssigem Stickstoff tiefgefroren. Wenngleich nicht als Einzelkolonien, sind diese Bibliotheken somit zumindest in Form von Agarplatten-Poolen dauerhaft für mögliche spätere Anwendungen gelagert. Die Poole dienten darüberhinaus in Analogie zu den Mikrotiterplatten-Poolen dem Screening. Kopien von Test-positiven Platten wurden gepickt, in Mikrotiterplatten überführt und wie oben beschrieben bis zur Identifikation positiver Einzelkolonien gescreent. Da die sehr zeitaufwendigen Arbeiten des Pickens und der „Poolbildung“ von Mikrotiterplatten nur für wenige Test-positive Agarplatten durchgeführt werden mußten, war es möglich, auch große Klonbibliotheken mit überschaubarem Zeitaufwand zu screenen.

2.7 PCR

2.7.1 PCR-Amplifikation

In dieser Arbeit wurde die Methode der Polymerase Ketten Reaktion (PCR, Mullis and Faloona, 1987) genutzt, um a) 16S rRNA-Gene sowie verschiedene funktionelle Gene (*pmaA*, *nifH*, *nifD* u.a.) mit spezifischen und universellen PCR-Assays aus Reinkulturen zu amplifizieren und b) große Klonbibliotheken nach Klonen zu durchsuchen, die genomische Fragmente der jeweiligen Zielgruppe beinhalten (Screening). Darüberhinaus wurde c) die Methode der "Reversen Transkription" angewendet, um mRNA in copy DNA (cDNA) umzuschreiben und diese anschließend mittels des RACE-Ansatzes (vgl. 2.10.7) zu charakterisieren.

Sämtliche PCR-Reaktionen wurden entweder im GeneAmp PCR System 9700 (Perkin Elmer Applied Biosystems, Weiterstadt) im Primus 96 Cyclyer (MWG Biotech, Ebersberg) oder im Eppendorf Gradientencycler (Eppendorf, Hamburg) durchgeführt.

Tabelle 2.2 : Standard-Reaktionsansatz für PCR-Amplifikationen

dNTPs (je 2 mM)	5 µl
10 x Taq-Puffer	5 µl
MgCl (25 mM)	3 µl
5'-Primer (30 pmol µl ⁻¹)	0,5 µl
3'-Primer (30 pmol µl ⁻¹)	0,5 µl
Taq-polymerase (5 U µl ⁻¹)	0,25 µl
Wasser(Sigma-Aldrich)	35 µl
Template	1 µl

Tabelle 2.3 : Standard-PCR-Profil zur Amplifikation von 16S rRNA und funktionellen Genen

Nr.	Temperatur	Dauer	Funktion
1	94° C	5 min	Denaturierung
2	Amplifikations-Zyklus		
2a	94° C	30 sec	Denaturierung
2b	54-68° C*	30 sec	Primer-Anlagerung
2c	72° C	1 min	DNA-Synthese
25-35 Wiederholungen von Schritt 2*			
3	72° C	7 min	DNA-Synthese
4	4°C	∞	Kühlprogramm

* Anlagerungstemperatur und Zyklenzahl variierten mit dem verwendeten Assay, vgl. Text.

2.7.2 Amplifikation von Nitrogenase-Genen aus Reinkulturen

Genomische DNA wurde wie unter 2.7.3 beschrieben extrahiert oder in bereits aufgereinigter Form von S. Dedysh zur Verfügung gestellt. Alle PCR-Reaktionsansätze beinhalteten 1,5 mM MgCl₂ und jeweils 330 nM der verwendeten

Primer (Vgl. Tab 2.2). Die PCR-Protokolle entsprachen dem in Tab. 2.3 angegebenen thermalen Profil mit den nachfolgenden Modifikationen.

Amplifikation des zentralen Bereichs des NifH-Gens:

Primer nifH-F1 und nifH-r, Anlagerungstemperatur 55° C, 35 Zyklen.

Amplifikation des zentralen Bereichs des NifD-Gens (1130 bp):

Primer nifD-f und nifD-r, Anlagerungstemperatur 59° C, 30 Zyklen.

Amplifikation des terminalen Bereichs des NifH-Gens, der intergenischen Region sowie des nahezu kompletten NifD-Gens (ca. 1900 bp):

Primer nifHD-F und nifD-r, Anlagerungstemperatur 63° C, Prozessierung (Elongation): 2 min, 35 Zyklen.

Im Rahmen dieser Arbeit wurden die Primer nifD-f, nifD-r und nifHD-f entwickelt, darüber hinaus wurden alle drei beschriebenen Assays mittels Gradienten-PCR anhand von Kontroll-DNA optimiert.

2.7.3 Extraktion von DNA aus Reinkulturen für die PCR-Amplifikation

50-250 ml Kultur wurden abzentrifugiert, in 1 ml kaltem (0-1°C) TPM-Puffer (Tris-HCl, pH 7,5; 50 mM Polyvinylpyrrolidon 25; 1,7% [w/v]; 20 mM MgCl₂) resuspendiert und 1 g Glasperlen ($\varnothing = 0,17 - 0,18$ mm, B. Braun International, Melsungen) zugegeben. Die Proben wurden 1 min bei 2400 upm im Zellhomogenisator (Mikrodismembrator, B. Braun International, Melsungen) geschüttelt und anschließend zentrifugiert. Sämtliche Zentrifugationsschritte wurden bei $14.600 \times g$ und 4°C durchgeführt. Der wäßrige Überstand (Überstand A) wurde zu 700 ml Phenol (kalt) gegeben und auf Eis gestellt. In einem zweiten Extraktionsschritt wurden 700 µl phenolgesättigten Lysis-Puffers auf den Niederschlag gegeben und wie oben verfahren. Der Überstand wurde zum Überstand A auf das Phenol gegeben, das Gemisch 1 min geschüttelt und 5 min abzentrifugiert. Der wäßrige Überstand wurde zu 700 ml Chloroform/Phenol [1:1 v/v] gegeben, 1 min ausgeschüttelt und 5 min abzentrifugiert. Der Überstand wurde in ein neues Gefäß überführt, mit einem Volumen Chloroform/Isoamylalkohol [24:1 v/v] versetzt, 1 min geschüttelt und danach 5 min zentrifugiert. Der wäßrige Überstand wurde in einem neuen Gefäß mit 0,1 Volumen 3 M Natrium-Acetat (pH 5,2) und 2 Volumen reinen Ethanol versetzt und gut durchmischt. Nach Inkubation (1 h, -80 °C) wurde die Probe 1 h bei 20° C

zentrifugiert. Der DNA-Niederschlag wurde mit 500 µl Ethanol (70%) gewaschen und für 10 min bei -20° C zentrifugiert, anschließend in der Vakuum-Zentrifuge getrocknet und in 100 µl TE-Puffer resuspendiert. Zu Kontrollzwecken wurden 10 µl auf ein Agarosegel aufgetragen, der Rest bei -20 °C gelagert.

2.7.4 Amplifikation von *pmoA1* und *pmoA2* aus Reinkulturen und Klonbibliotheken

Die Amplifikation der zentralen Bereiche von *pmoA1* und *pmoA2* aus Reinkulturen erfolgte mittels zweier PCR-Assays, wobei einer bei 66° C Anlagerungstemperatur spezifisch für *pmoA2* ist, bei 60° C jedoch auch zur Amplifikation von konventionellen *pmoA*-Genen (*pmoA1*) genutzt werden kann. Alle PCR Reaktionsansätze enthielten 1,5 mM MgCl₂ und jeweils 125 nM der verwendeten Primer (vgl. Tab. 2.2). Das thermale Profil folgte der Tabelle 2.3.

Amplifikation des zentralen Bereichs der konventionellen *pmoA1*:

A189f und A682b, Anlagerungstemperatur 50° C, 30 Zyklen.

oder: A206f und A703b, Anlagerungstemperatur 60° C, 30 Zyklen.

Amplifikation des zentralen Bereichs der „*novel*“-*pmoA2*:

A206f und A703b, Anlagerungstemperatur 66° C, 30 Zyklen.

2.7.5 Amplifikation von *pmoA*-Sequenzen des USCα aus Umweltproben und Klonen

Für die spezifische Amplifikation des zentralen Bereichs (507 bp) der *pmoA*-Sequenzen des USCα wurde ein von Kolb und Mitarbeitern (2004) entwickeltes Assay verwendet. Klone sowie Screening-Poole wurden durch aufkochen (10 min 100°C) lysiert, 5 min bei 6000 g zentrifugiert und der klare Überstand als Template eingesetzt. Alternativ wurde aus der Umwelt extrahierte DNA in verschiedenen Verdünnungen verwendet. Die PCR-Ansätze beinhalteten 1,5 mM MgCl₂ und jeweils 1000 nM der Primer A189f und USCα675b (vgl. Tab 2.2). Das thermale Profil folgte der Tabelle 2.3. Die Anlagerungstemperatur betrug 67°C.

Um die relative Aufschluß-Effizienz verschiedener DNA-Extraktions-Methoden hinsichtlich der Zielgruppe zu verfolgen, wurde die extrahierte Umwelt-DNA verdünnt (10^{-1} bis 10^{-5}), und die so erhaltenen Verdünnungsreihen mit dem USC α -spezifischen Assay sowie mit einem für alle *Bacteria* universellen 16S rDNA-Assay amplifiziert (Daten nicht gezeigt).

2.7.6 PCR basiertes Screening

Für das Screening der Klonbibliotheken nach Klonen, die DNA-Fragmente der jeweiligen Zielgruppe aufgenommen hatten, wurden die nachfolgend beschriebenen Assays verwendet. Das Screening der *Methylocystis* sp. Stamm SC2-Bibliothek wie auch der *M. acidiphila*-Bibliothek nach Kopien der *pmoA1* wurde mit dem Assay "189f / 682b" (vgl. 2.7.4) durchgeführt. Das Screening nach USC α -Fragmenten in der Umwelt-Bibliothek wurde sowohl mit diesem Assay, als auch mit dem für USC α -Sequenzen spezifischen Assay 189f/For625b (Kolb et al., 2003) durchgeführt. Als Positivkontrollen dienten Agarplatten- bzw. Mikrotiterplatten-Poole (vgl. 2.6.9) die mit stark verdünnten Extrakten genomischer DNA von *Methylocystis* sp. Stamm SC2 vermischt wurden. Für das USC α -spezifische Screening wurde DNA eines USC α -Umweltklons (bereitgestellt von S. Kolb) anstelle der genomischen DNA verwendet.

2.8 Analytische Agarose-Gelelektrophorese

Die Agarose-Gelelektrophorese ist eine klassische Methode zur Auftrennung von Nukleinsäuren nach ihrer Größe. Negativ geladene Moleküle, aufgrund der Phosphatgruppen auch DNA, bewegen sich im elektrischen Feld auf die Anode der Elektrophoresekammer zu. Ihre Geschwindigkeit ist dabei proportional zur Ladungsdichte, Größe und Form (Voet et al., 1998). Agarose bildet ein lockeres Geflecht aus Kohlenhydratpolymeren, das die DNA in Abhängigkeit von ihrer Größe in ihrer Bewegung hemmt. Im Falle linearisierter DNA ist die Wanderungsgeschwindigkeit dem Logarithmus ihres Molekulargewichts umgekehrt proportional.

Nach der Auftrennung der DNA wurden die Gele mit Ethidium-Bromid (0,01%ig) gefärbt und bei UV-Licht der Wellenlänge 302 nm aufgenommen. Durch Vergleich mit flankierend aufgetragenen DNA-Standards (Smart Ladder, Eurogentech) konnten Größe und Konzentration der einzelnen Fragmente bestimmt werden. Für diese Arbeit wurden je nach aufzutrennenden Fragmentgrößen 0,6 bis 1,5%ige Agarosegel verwendet. Je nach Konzentration wurden unterschiedliche Mengen Probe mit Wasser auf 9 µl aufgefüllt, mit 1 µl Auftragspuffer gemischt und aufgetragen. Die Auftrennung erfolgte bei ca. 6V cm⁻¹ für 30 bis 90 min in 1x TAE-Puffer (40 mM Tris/Acetat, pH 8; 1 mM EDTA)

2.9 Sequenzierungen

Die Sequenzierung von PCR-Produkten und aufgereinigten Plasmiden („Direktsequenzierung“) erfolgte nach der von Sanger entwickelten Kettenabbruch-Methode (Sanger *et al.*, 1977) mittels des „ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit“ (PE Applied Biosystems, Weiterstadt). Bei dieser Methode werden neben unmarkierten Desoxynukleotiden Fluoreszenz-markierte Didesoxynukleotide in eine zyklische Reaktion eingesetzt. Im Unterschied zur PCR wird lediglich ein Primer zugegeben. Wird während der Polymerase-Reaktion ein Didesoxynukleotid eingebaut, kann der neu-synthetisierte Strang nicht verlängert werden, es kommt zum „Kettenabbruch“. Die so entstehenden Fragmente unterschiedlicher Größe sind je nach Abbruch-Nukleotid mit einem anderen Fluoreszenz-Farbstoff markiert. Nach größenspezifischer Auftrennung in einem denaturierenden Polyacrylamidgel werden die Fluoreszenz-Farbstoffe angeregt, die spezifischen Emissionen noch während der Elektrophorese gemessen und anschließend in die entsprechende Basenabfolge ("Sequenz") übersetzt.

2.9.1 Sequenzierung von PCR-Produkten

Vor der Sequenzierung wurden PCR-Produkte mit dem QIAquick PCR-Purifikation Kit™ aufgereinigt und die DNA-Menge durch gelelektrophoretische Auftrennung und Vergleich mit Standards bekannter Quantität abgeschätzt. Die Reaktionsansätze

enthielten 50-100 ng DNA, 6 pmol Sequenzierungs-Primer und 3 µl Terminator Ready Reaction Mix (PE Applied Biosystems) in einem Gesamtvolumen von 10 µl. Nach 25 Zyklen (96° C für 10 s, 50° C für 5 s und 60° C für 4 min) wurden unverbrauchte Primer und Ribonukleinsäuren mittels Microspin G-50 Chromatographiesäulen (Pharmacia Biotech) entfernt, die DNA unter Vakuum getrocknet und in 2 µl Formamide Loading Dye (Amersham Biosciences) aufgenommen. Vor dem Beladen des Sequenziergels wurde die DNA für 5 min bei 100° C denaturiert. Die anschließende Lagerung auf Eis verhinderte die Renaturierung der PCR-Produkte. Die elektrophoretische Auftrennung der Fluoreszenz-markierten Einzelstränge wurde am ABI PRISM 377 bzw. am ABI PRISM 330 DNA Sequenzer (PE Applied Biosystems) durchgeführt. Die Auswertung der Rohdaten ("Base calling" etc.) erfolgte mittels der zugehörigen ABI-Software.

2.9.2 Direktsequenzierung ohne vorhergehende PCR -Amplifikation

Im Gegensatz zu der mit zwei Primern durchgeführten PCR steigt die Menge der in der Reaktion enthaltenen DNA bei der zyklischen Sequenzierung nicht exponentiell, sondern nur linear an. Um hinreichende Mengen Fluoreszenz-markierter Polynukleotide zu erzeugen, wird deshalb in der Regel die zu sequenzierende DNA mittels vorheriger PCR amplifiziert (vgl. 2.9.1). Soll jedoch ein Sequenzabschnitt bestimmt werden, für den nur eine Primerbindestelle bekannt ist, ist die Durchführung der PCR nicht möglich. Im Verlauf dieser Arbeit trat dieses Problem auf, wenn a) Randbereiche von großen Insert sequenziert werden sollten („End-Sequenzierung“) oder b) flankierenden genomischen Bereiche eines bekannten Sequenz-Abschnitts gelesen werden sollten (beispielsweise während der Sequenzierung der Operon-Struktur der pMMO-1 von *Methylocystis* sp. Stamm SC2 ohne vorhergehende „Shotgut-Subsequenzierung“, vgl. 2.9.3). Im Fall a) wurden die Vektor-spezifischen Primer pCCFosF und pCCFosR (Tab. 2.4) verwendet. Für die unter b) beschriebene Sequenzierung flankierender genomischer Bereiche wurde die Technik des „Primer-Walkings“ etabliert, die nachfolgend am selben Beispiel erläutert wird. Ausgehend von der mit universellen Primern amplifizierten zentralen Region der *pmoA1* wurden Primer formuliert, mit deren Hilfe die angrenzenden Bereiche sequenziert werden konnten. Die so gewonnenen Sequenz-Informationen dienten der Erstellung neuer Primer. Der Zyklus wurde fortgeführt, bis die vollständige Sequenz des Operons

sowie jeweils ca. 1000 Basen der flankierenden Sequenzen gelesen waren. Basierend auf den erhobenen Sequenz-Daten wurden weitere Primer formuliert, die die Sequenzierung der Gegenstränge erlaubten, was das Erreichen hinreichender Sequenzier-Genauigkeit ermöglichte (vgl. Abb 2.6).

Da im Gegensatz zur Sequenzierung von PCR-Produkten bei der Sequenzierung ohne vorhergehende PCR-Amplifikation nur eine geringe Menge Template eingesetzt wird, wurde der Reaktionsansatz und das thermale Profil der Zyklischen Sequenzierung wie folgt abgeändert. Jede Reaktion enthielt 13 µl aufgereinigte DNA, 6 pmol Sequenzierungsprimer und 6 µl "Terminator Ready Reaction Mix" (PE Applied Biosystems) in einem Gesamtvolumen von 20 µl. Das zyklische Profil (96°C für 30 s, 50°C für 20 s und 60°C für 4 min) wurde 99 mal durchlaufen. Die Produkte wurden wie beschrieben (vgl. 2.9.1) aufgereinigt und sequenziert.

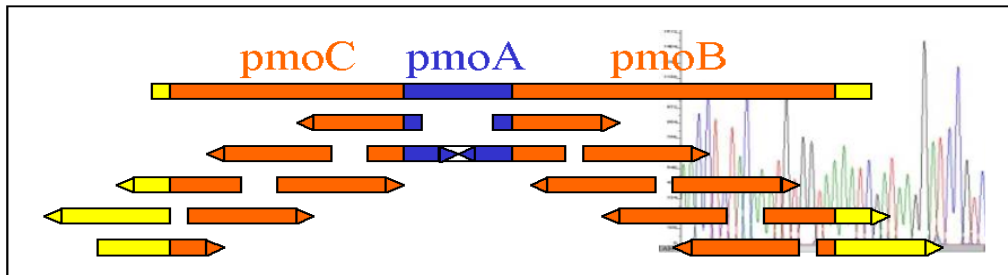


Abb. 2.6 : Prinzip der "Primer-Walking"-Sequenzier-Strategie. Sequenzierung des *pmoCAB1*-Operons von *Methylocystis* sp. Stamm SC2 ohne vorhergehende PCR-Amplifikation oder Subklonierung ausgehend von der bekannten Sequenz des zentralen Bereiches von *pmoA1* (blau dargestellt).

2.9.3 "Shotgun"-Sequenzierung genomischer Fragmente

Die Kompletsequenzierung der "Vektor/Insert-Konstrukte" Test-positiver Klone wurde am MPI für molekulare Genetik, Berlin von M. Kube und Mitarbeitern durchgeführt. Die Large Insert-Plasmide (BACs bzw. Fosmide) wurden extrahiert, die DNA mittels Ultraschall fragmentiert (durchschnittliche Fragmentgröße 1,5 kb), größenselektiert, mit einem pUC19 Vektor (Fermentas, St. Leon-Rot) ligiert und via Elektroporation in *E. coli* (Stamm DH10B; Gibco, Karlsruhe) transformiert ("Subklonierung"). Die Inserts der so erstellten Subklone wurden amplifiziert und die PCR-Produkte mit Dye Terminator Chemistry sequenziert (3730 ABI Capillary

Sequencer Systems, ABI, Weiterstadt). Bereiche geringerer Qualität wurden gezielt resequenziert, nicht vereinigte Contigs mittels "Primer-Walking" verbunden. Die finalen Sequenzen wurden als "Phred 30" qualifiziert, was einer Sequenziergenauigkeit von 99,9 % und einer durchschnittlich sechs- bis achtfachen Sequenz-Abdeckung entspricht.

2.10 Bestimmung von Transkriptions-Startpunkten

2.10.1 Arbeiten mit RNA

Aufgrund der Anfälligkeit von RNA für ubiquitär vorkommende RNasen werden RNA-Arbeiten unter Beachtung besonderer Regeln durchgeführt. Alle Arbeiten wurden mit Latexhandschuhen auf einem ausschließlich für RNA-Arbeiten genutzten, regelmäßig mit "RNaseAway" (Invitrogen, Karlsruhe) behandelten Arbeitsplatz durchgeführt. RNaseAway inhibiert laut Hersteller RNase-Aktivität zu 100%. Soweit möglich, wurden zertifiziert RNase-freie Chemikalien verwendet. Plastikwaren wurden, soweit nicht vom Hersteller als RNase-frei gekennzeichnet, mit RNaseAway behandelt, Glaswaren bei 240°C 6h ausgebacken. Verwendete Lösungen wurden mit Diethyl-Pyrocarnat- (DEPC) Wasser angesetzt. Alle nicht ausdrücklich ausgenommenen Arbeitsschritte erfolgten auf Eis.

2.10.2 Extraktion von Gesamt-RNA mit dem RNeasy (Mini)–Kit

Etwa 1×10^9 Zellen *Methylocystis sp.* Stamm SC2 wurde für 20 min bei 5000 g abzentrifugiert, in 2 ml Medium resuspendiert und mit zwei Volumen RNA-Protect (Qiagen, Hilden) versetzt, um die RNA zu stabilisieren. Nach 5 min Inkubation wurden die Zellen abzentrifugiert und das Pellet anschließend in 100 µl RNA-Protect resuspendiert. Die Zellsuspension wurde anschließend mit 0,5 g Glaskugeln (ø 0,17-0,18 mm; B. Braun) für 1 min bei 2500 upm im "Beadbeater" („Micro Dismembrator S“, B. Braun) aufgeschlagen. Beads und Zelltrümmer wurden bei 7000 g und 4° C für 30 s abzentrifugiert, der Überstand abgenommen und nach einer modifizierten Version des "RNeasy Mini Kit" - Protokolls behandelt.

Nach Zugabe von 350µl Puffer RLT und starkem Durchmischen (Vortexen) wurden unlösliche Bestandteile für 2 min bei 10.000 g abzentrifugiert, der Überstand anschließend mit 200µl 96% EtOH versetzt, durch Pipetieren vermischt und auf eine "RNeasy-Mini-Säule" in einem 2 ml Eppendorf-Auffanggefäß gegeben. Nach Zentrifugation für 15 s bei 10000 g in der Mikrozentrifuge wurde der Durchlauf verworfen, und die Säule zweimal mit 750µl Puffer RW1 gewaschen (Zentrifugation für 15 s bei 10.000 g). Anschließend wurde die Säule in ein steriles 2 ml Eppendorfcup überführt, mit 500µl Puffer RPE wie oben gewaschen und der Durchlauf erneut verworfen. Nach erneutem Waschen mit 500µl Puffer RPE wurde die Silikamembran der Säule durch 2 minütiges Zentrifugieren bei 13.000 g getrocknet. Abschließend wurde die Säule erneut in ein sauberes 2 ml Eppendorfcup überführt und die RNA mit 50 µl RNase-freiem Wasser eluiert. Das Wasser wurde direkt auf die Membran appliziert, im Anschluß wurde für 1 min bei 13.000 g zentrifugiert. Um die Extraktionseffizienz zu erhöhen, wurde der letzte Schritt noch einmal wiederholt und die Eluate vereinigt. Die RNA wurde anschließend photometrisch quantifiziert (Messung der E_{260}). 5µl wurden auf einem Agarosegel aufgetrennt um die Qualität der Extraktion zu überprüfen. Intakte RNA bildet im Gel zwei distinkte Banden, die der 16 und 23S rRNA entsprechen. Wird die RNA dagegen durch RNasen teilweise abgebaut, sind keine diskreten Banden sichtbar.

2.10.4 Extraktion von Gesamt-RNA mittels Trizol-Reagent

Alternativ zur Extraktion von Gesamt-RNA mit dem RNeasy (Mini)-Kit (2.10.3) wurde RNA auch mittels Trizol-Reagent (Invitrogen) isoliert. Trizol ist eine einphasige Lösung aus Phenol und Guanidin-Isothiocyanat. Es ermöglicht die RNA-Extraktion in einem Schritt. Ca. 1×10^9 Zellen wurden in 8 ml Trizol resuspendiert und die Lösung auf 2 ml Eppendorfgefäße verteilt. Anschließend wurden die Zellen im Beadbeater mit 0,5 g Glasperlen lysiert und die Zelltrümmer abzentrifugiert. Der Überstand wurde 5 min bei Raumtemperatur inkubiert und mit 1/5 Volumen Chloroform versetzt. Das Gemisch wurde anschließend für 15 geschüttelt und für weitere 3 min inkubiert. Die Trennung von wässriger und organischer Phase erfolgte durch Zentrifugation bei 10.000 g für 15 min bei 4°C. Die RNA befand sich in der farblosen wässrigen Phase. Diese wurde vorsichtig abgenommen und mit 4 ml Trizol Reagenz versetzt. Nach weiteren 15 Minuten Inkubation wurde die RNA für 15 min

bei 10.000 g und 4°C präzipitiert. Die RNA bildete dabei ein durchscheinendes Pellet. Dieses wurde mit 1 ml kaltem 75%igem Ethanol gewaschen und anschließend 10 min an der Luft getrocknet. Die RNA wurde in 40 µl DEPC-Wasser resuspendiert und anschließend per Agarose-Gelelektrophorese sowie photometrisch quantifiziert.

2.10.5 Abbau der DNA mittels DNase

Um möglicherweise die RNA kontaminierende DNA zu entfernen wurde im Anschluß an die RNA-Extraktion eine DNase-Behandlung mit RQ1 RNase-Free DNase (Promega) durchgeführt. Dabei wurde 5 U DNase pro µg RNA eingesetzt. Die Reaktion wurde 45 min bei 37°C inkubiert. Um die DNase zu deaktivieren, wurde 1/10 Volumen Stop Solution (Promega) zugegeben und für 10 min bei 65°C inkubiert. Der vollständige Abbau der DNA wurde durch PCR-Reaktionen mit *pmoA*-spezifischen Primern überprüft (Tab. 2.5). Nur wenn kein PCR-Produkt generiert werden konnte, wurde die RNA für die cDNA-Synthese eingesetzt.

2.10.6 cDNA-Synthese

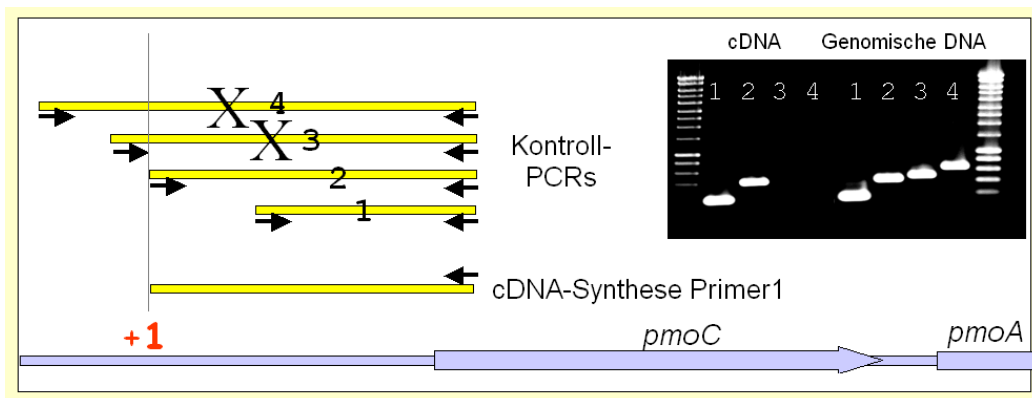
Mit dem Begriff "RT-PCR" (Reverse-Transkriptions-PCR) wird das Umschreiben von mRNA in cDNA („copy“ oder komplementäre DNA) durch das Enzym Reverse-Transkriptase bezeichnet. Für die RT-PCR wurde das ThermoScript reverse transcription-PCR System (Invitrogen) nach Angaben des Herstellers verwendet. Für die spezifische reverse Transkription der konventionellen mRNA wurde der Primer ConvC100r (spezifisch für *pmoC1*), für die mRNA des novel-Operons der Primer NovC313r (spezifisch für *pmoC2*) entwickelt. Die cDNA-Synthese erfolgte nach dem Protokoll des Herstellers für 1 h bei 56°C. Der Ansatz ist in Tab. 2.4 dargestellt. Im Anschluß wurde die cDNA mit dem PCR-Purification Kit (QIAGEN) aufgereinigt.

Tab. 2.4 : Reaktionsansatz für die "Reverse-Transkriptions-PCR"

Komponente	Volumen [μl]
Primer, $17 \text{ pmol } \mu\text{l}^{-1}$	1,0
RNA (abs. $0,1 \text{ } \mu\text{g}$)	x
dNTPs, 10 mM	2,0
5x cDNA-Synthesepuffer	4,0
DTT, $0,1 \text{ M}$	1,0
RNaseOUT, $40 \text{ U } \mu\text{l}^{-1}$ (Invitrogen)	1,0
ThermoScript TM RT, $15 \text{ U } \mu\text{l}^{-1}$	1,0
DEPC Wasser	auf $20 \text{ } \mu\text{l}$ auffüllen

2.10.7 Bestimmung der Promoter-Regionen mittels RACE

Die Transkriptions-Startpunkte von *pmoCAB1* und *pmoCAB2* von *Methylocystis* sp. Stamm SC2 wurden mittels der "Rapid Amplification of cDNA Ends" (RACE) – Technik ermittelt. Dem RACE-Ansatz liegt folgende Idee zu Grunde: Das 3'-Ende der mittels RT-PCR (2.10.6) synthetisierten cDNA entspricht dem 5'-Ende der transkribierten mRNA. Kann dieser cDNA-Terminus ermittelt werden, so ist auch der Transkriptions-Startpunkt bekannt. Um die nur in geringen Mengen vorliegende cDNA mittels PCR zu vervielfältigen, wird zunächst das 3'-Ende mittels des Enzyms Terminale Transferase polyadenyliert. Dieses Poly(A)-Ende dient nun als Primer-Bindestelle für einen zusammengesetzten Primer, dessen 3'-Terminus aus 16 Thymidinen besteht. In Kombination mit einem cDNA spezifischen Gegenprimer kann nun die cDNA amplifiziert werden. Um eine höhere Spezifität zu erreichen, wurde in einer zweiten PCR-Reaktion eine 1:50 Verdünnung des ersten PCR-Produkt erneut amplifiziert. Die Spezifität wird durch Verwendung eines weiteren cDNA-spezifischen internen Primers ermöglicht, der Gegenprimer bindet an die dem 5'-Terminus des zusammengesetzten Primers komplementäre Region. Das so erzeugte PCR-Produkt wird aufgereinigt und sequenziert. Der Beginn der Poly(A)/Poly(T)-Region entspricht dem letzten nicht transkribierten Nukleotid des Operons. Das Prinzip der RACE sowie die verwendeten Primer sind in Abb. 2.7 dargestellt, die Primersequenzen in Tab 2.5. In dieser Arbeit wurde das 5'/3' RACE Kit (Roche, Mannheim) gemäß den Angaben des Herstellers verwendet.



2.11 Bioinformatik

2.11.1 Datenbanken

Für die Phylogenetischen Analysen, die im Rahmen dieser Arbeit durchgeführt wurden, wurde hauptsächlich das Programm-Paket „ARB“ (Ludwig et al., 2004) verwendet. Kern von ARB ist eine Graphikorientierte Datenbank-Verwaltung für biologische Sequenzen und zugehörige Meta-Daten. Daran angegliedert sind Werkzeuge für die Formulierung von Primern oder Sonden, automatische und manuelle Alignment-Programme sowie eine Vielzahl Phylogenetischer Algorithmen. Neben einer umfassenden 16S/18S rDNA-Datenbank mit mehr als 48.000 Einträgen (Stand 2004) verfügt ARB auch über die Möglichkeit individuelle Datenbanken, beispielsweise für funktionelle Gene, zu erzeugen und zu verwalten. Datenbanken folgender Gene/Proteine wurden für diese Arbeit erstellt: *pmoA/amoA* (PmoA/AmoA), *pmoCAB/amoCAB* (PmoCAB/AmoCAB), *nifD*(NifD), *nifH*(NifH), Staphylococcem-ähnliche Nuklease, Zink-Metallo-Protease, Prephenate-Dehydratase, CowB sowie Datenbanken diverser, hier nicht weiter diskutierter Gene.

Vereinfachend wurden dazu folgende Arbeitsschritte durchgeführt:

- 1) Suche öffentlich zugänglicher Sequenzen mittels BLAST/FASTA bzw. Entrez/SRS in Genbank bzw. anderen öffentliche Datenbanken (vgl.2.11.11)
- 2) Anlegen einer neuen ARB-Datenbank, Import der Sequenz- und Meta-Informationen
- 3) Übersetzen von DNA Sequenzen in Polypeptide
- 4) Automatische Erstellung eines multiplen Alignments der abgeleiteten Polypeptide mittels des Algorithmus ClustalW, manuelle Kontrolle und Korrektur des Alignments.
- 5) Alignment der Nukleinsäuresequenzen anhand der Aminosäure-Alignments.
- 6) Erstellen und Einfügen weiterer, insbesondere proprietärer Sequenzen

Die Sequenzen der *pmoCAB/amoCAB* (PmoCAB/AmoCAB) Datenbank wurden durch Konkatenieren der einzelnen Gene/Polypeptide unter Auslassung intergenischer Regionen manuell konstruiert. Alternativ zu den Schritten 2-5 wurden bereits alignte Sequenzen von entsprechenden öffentlichen Datenbanken (PFAM, Protein families

database of alignments and HMMs, <http://www.sanger.ac.uk/Software/Pfam/>; COG Clusters of Orthologous Groups of Proteins, <http://www.ncbi.nlm.nih.gov/COG/>) bezogen und in die neu erstellte ARB-Datenbank übernommen.

2.11.2 Phylogenetische Analysen

Zwei Strategien der sequenzbasierten phylogenetischen Rekonstruktion sind prinzipiell zu unterscheiden. Bei distanzbasierten Algorithmen (hier verwendet: Neighbor Joining Methode (NJ; Saitou und Nei, 1987)) wird in einem ersten Schritt die zu verrechnende Information radikal reduziert. Auf Basis aller in die Kalkulation einzubeziehender Sequenzen wird eine sogenannte „Distanz-Matrix“ berechnet, in der die prozentualen Abweichungen ("Unähnlichkeitswerte") aller Paare von Sequenzen festgehalten werden. In der Regel werden diese Werte mit Hilfe eines Korrekturfaktors (Evolutionsmodell, z.B.: Kimura (DNA, Kimura, 1980) JK (DNA, Jukes und Cantor, 1969); JTT (Proteine, Jones et al., 1992) extrapoliert, um in der berechneten Distanz („Hamming-Distanz“) nicht enthaltene Mutationsereignisse (beispielsweise zwei aufeinanderfolgende Mutationen an einer Position) oder Homoplasien zu berücksichtigen. Ebenfalls üblich ist der Ausschluß (Maskierung) hochvariabler Sequenzabschnitte, da hohe Mutationsraten in nahezu zufällig verteilten Nukleotid-/Aminosäure-Frequenzen resultieren die keine brauchbare phylogenetische Informationen mehr enthalten (sondern im Gegenteil die Berechnung negativ beeinflussen könnten). Distanzbasierte Phylogenien (beispielsweise NJ-, UPGMA-Methode) beruhen ausschließlich auf den in der Distanz-Matrix gespeicherten paarweisen Distanzen. Im Gegensatz dazu werden bei positionsspezifischen Algorithmen (In dieser Arbeit verwendet: Maximum Likelihood (ML)-Methoden (TreePuzzle, PROTml, DNaml) sowie Maximum Parsimony (MP)-Methoden) alle Sequenzinformationen berücksichtigt. Dabei werden alle Positionen des Alignments (nach Maskierung) unabhängig voneinander verrechnet. Im Gegensatz zu den distanzbasierten Methoden wird nicht nur die (Nicht-) Übereinstimmung zweier Reste in die Kalkulation mit einbezogen, sondern auch die Wahrscheinlichkeit mit der zwei homologe, möglicherweise von einander abweichende Reste während der Evolution aus einem gemeinsamen Vorläufer hervorgegangen sein könnten. Obwohl mehr Information verrechnet wird, müssen positionsspezifische Algorithmen nicht immer in realistischeren Phylogenien resultieren. Parsimony-Methoden beispielsweise sind

aufgrund restriktiver Modellannahmen („Minimum Evolution-Prinzip“) für die Verrechnung von molekularen Sequenzen nur bedingt anwendbar. Eine Übersicht über phylogenetische Algorithmen ist beispielsweise bei Swofford und Olsen (1990) zu finden.

Allen in dieser Arbeit gezeigten Stammbäumen liegt eine Vielzahl phylogenetischer Berechnungen zugrunde. So wurden immer verschiedene Algorithmen (NJ, ML) in Kombination mit mehreren Evolutionsmodellen (JC, PAM, JTT, WAG) genutzt. Soweit die Menge zu verrechnender Daten dies zuließ und die Fragestellung es erforderte, wurde die Auswahl zu verrechnender Spezies leicht variiert, um den Einfluß möglicher zufälliger Sequenz-Übereinstimmungen zu minimieren. Aus ähnlichen Gründen wurden bei vielen distanzbasierten, aber auch bei Likelihood-basierten* Berechnungen kombinierte Bootstrap-Analysen durchgeführt. Hierbei wird der (maskierte) Primärdatensatz durch zufällige Gewichtungen manipuliert. Einzelne Positionen des Alignments können einmal, mehrfach oder auch gar nicht in eine Berechnung einfließen. Der Bootstrap-Algorithmus erzeugt so eine Vielzahl bewusst verfälschter Originaldatensätze (in dieser Arbeit beispielsweise 1000 Varianten für NJ-Analysen) von denen jeder als Grundlage einer phylogenetischen Berechnung genutzt wird. Die erzeugten phylogenetischen Bäume werden anschließend automatisch miteinander verglichen und verrechnet. Für jeden Ast des so erzeugten Konsensus-Baums wird die relative Häufigkeit angegeben, mit der die hier gemeinsam verzweigenden Organismen (Sequenzen) in den verschiedenen Bäumen zusammen gruppierten.

Die Ergebnisse der einzelnen phylogenetischen Berechnungen wurden anschließend miteinander verglichen. Nicht eindeutig aufzulösende Bereiche wurden manuell in Multifurkationen überführt. Im Gegensatz zu bifurkativen Verzweigungen, die implizit die Monophylie jeweils zweier gemeinsam verzweigender Linien postulieren, und somit jeder Linie eine nächst-verwandte Linie zuordnen, ist bei

* Eine Bootstrap-Analyse wurde nicht in Kombination mit dem Tree-Puzzle-Algorithmus durchgeführt. Bei dieser Berechnungsmethode werden zunächst alle möglichen Kombinationen von jeweils 4 Sequenzen eines Datensatzes miteinander verrechnet. Die entstehenden Quartetbäume gruppieren jeweils 2 der 4 Sequenzen miteinander. In einem zweiten Schritt werden nun die in den Quartetbäumen vorliegenden Informationen genutzt, um einen Stammbaum aller verrechneter Sequenzen zu erzeugen. Da die Reihenfolge der Inkorporation der Quartet-Bäume in die laufende Berechnung einen drastischen Einfluß auf das Ergebnis hat, wird eine Vielzahl (in dieser Arbeit 10.000-25.000) Bäume berechnet wobei die Reihenfolge der Sequenzinkorporation zufallsgemäß variiert wird. Bei der Berechnung des Konsensusbaums werden automatisch Angaben zu der relativen Häufigkeit aller Verzweigungen (ähnlich den Bootstrap-Werten) gemacht.

Multifurkationen keine Auflösung der internen Relationen der gemeinsam verzweigenden Linien gegeben.

2.11.3 Computergestützte Sekundärstruktur-Analysen von PmoCAB

Für die Erstellung von Hydrophobizitätsplots bzw. die Berechnung membranspannender Regionen der jeweiligen abgeleiteten Polypeptide der pMMO wurden verschiedene Programme benutzt und die Ergebnisse miteinander verglichen, u.a.:

- 1) "Protein Hydrophilicity/Hydrophobicity Search and Comparison Server" (<http://bioinformatics.weizmann.ac.il/hydroph/>), Weizmann Institute of Science, Israel.
- 2) toppred – transmembrane topology prediction program (<http://bioweb.pasteur.fr/seqanal/interfaces/toppred.html>). (Claros and von Heijne, 1994)
- 3) "Trans membrane Hidden Markov Model -TMHMM tool, Swiss Institute of Bioinformatics , <http://www.expasy.ch/tools/#transmem>

Das prinzipielle Vorgehen der beiden erstgenannten Programme ist wie folgt (vereinfacht): Jeder Aminosäuren wird aufgrund ihrer physiko-chemischen Eigenschaften ein numerischer Wert zugeordnet (Beispielsweise nach Kyte und Doolittle (1982), von Heijne (1992) oder Hopp und Woods (1978)), der den Grad der Hydrophobizität der Aminosäure reflektiert. Hydrophobizitätsplots basieren auf solchen Hydrophobizitäts-Indizes der Aminosäuren und stellen die Hydrophobizität als Funktion der Aminosäuresequenz einer Polypeptidkette dar. Dazu werden die Werte von beispielsweise jeweils 11 benachbarten Aminosäuren gemittelt und die Werte graphisch dargestellt. Unter Berücksichtigung weiterer Modellannahmen (z.B. „positive inside rule“; von Heijne, 1992) können im Anschluß Vorhersagen über die Sekundärstruktur des analysierten Proteins getroffen, bzw. membranspannende Regionen bestimmt werden.

Programme wie das unter 3) genannte TMHMM basieren auf Sequenzstatistiken. Als Grundlage dienen sogenannte "versteckte Markov-Modelle" (HMMs). Vereinfachend dargestellt, wird ein Datensatz bekannter Transmembran-Helices sowie ein Satz zufälliger Sequenzen genutzt, um das Modell zu "trainieren".

Das Training besteht aus einer statistischen Analyse der Trainings-Daten. Das Modell "kennt" nun die Wahrscheinlichkeiten, mit denen in Transmembran-Helices (bzw. Zufalls-Sequenzen) bestimmte Aminosäuren (in einer bestimmten Abfolge) vorkommen ("emittiert werden"). Mittels rekursiver Algorithmen kann nun für eine gegebene Sequenz berechnet werden, wie groß für eine gegebene Region die Wahrscheinlichkeit ist, daß sie eine Transmembran-Helix formt bzw.: Transmembran-Helices werden so vorhergesagt, daß die Wahrscheinlichkeit der Emission der realen Aminosäure-Abfolge maximiert wird. HMMs werden in vielen Bereichen der Bioinformatik verwendet, beispielsweise auch bei der Annotation genomischer Fragmente ("ORF-Prediction", 2.11.9).

2.11.4 Identifikation konservierter Aminosäure-Reste

Grundlage für die Identifikation hochkonservierter Sequenzbereiche ist die Zuordnung homologer Aminosäuren bzw. Nukleotide („Alignment“). Tukvathullin und Mitarbeiter (2000) identifizierten auf Basis aller zu diesem Zeitpunkt öffentlich zugänglichen PmoC-, PmoA-, PmoB-, AmoC-, AmoA- und AmoB-Sequenzen einen Satz hochkonservierter Aminosäuren, die jeweils in nahezu allen homologen Sequenzen vorhanden waren. Der Vergleich der in dieser Arbeit analysierten Operons mit der hochkonservierten Konsensussequenz basiert auf den unter 2.11.1 beschriebenen multiplen Alignments sowie auf paarweise globalen Alignments. Letztere wurden mit dem Programm „Stretcher“ (vgl. 2.11.11) unter Verwendung der BLOSUM-Substitutionsmatrizen-Familie erstellt. Die abschließende Analyse erfolgte manuell.

2.11.5 Signaturanalyse von PmoA1, AmoA und PmoA2

Die im Zuge der funktionellen Charakterisierung der pMMO-2 durchgeführte Analyse spezifischer Signaturen von PmoA1, PmoA2 und AmoA wurde mit Hilfe des Programms ARB (vgl. 2.11.1) durchgeführt. Ein Alignment von zum Teil partiellen 919 AmoA-, 315 PmoA1- und 40 PmoA2-Sequenzen (vgl. 2.11.1) diente als Grundlage für die Ermittlung gruppenspezifischer Signaturen. Dazu wurden mit ARB-implementierten Werkzeugen Konsensussequenzen der einzelnen Gruppen mit jeweils a) 90% und b) 97% minimaler Übereinstimmung berechnet. D.h.: In die

Konsensussequenzen gingen nur solche Reste ein, die in 90% (97%) der Sequenzen der jeweiligen Gruppe vorhanden sind. Daneben wurden in analoger Weise Signaturen auf Basis funktionell ähnlicher Aminosäure-Gruppen (Gruppen im 1-Buchstaben-Kode: PAGST, QNEBZ, HKR, LIVM, FYW) berechnet. Hierbei wurden auch solche Positionen in die Signatur aufgenommen, an denen mehr als 90% (97%) der verglichenen Sequenzen eine Aminosäure der selben Gruppe aufweisen. Durch Vergleich der Signaturen von PmoA1 und AmoA konnten 15 (18) Positionen identifiziert werden, an denen die beide Proteine unterschiedliche Aminosäure-(Gruppen-) Signaturen aufweisen, und die somit potentiell eine Unterscheidung ermöglichen. Anschließend wurde für die Signaturen der PmoA2-Sequenzen überprüft, ob sie mit den für PmoA1 bzw. AmoA charakteristischen Resten (Gruppen) übereinstimmen.

2.11.6 Entwicklung spezifischer Primer für die PCR

Mit Hilfe der Programme Probe-Design und Probe-Match der ARB-Computersoftware (vgl. 2.11.1) wurden folgende Primer bzw. Sonden für die spezifische Amplifikation funktioneller Gene formuliert: *nifH*Df, *nifD*f, *nifD*r, PRO*pmoA2* (Tab. 2.5). Der Vergleich der in der jeweiligen Datenbank verfügbaren Sequenzen ermöglichte die Definition von für *nifH* bzw. *nifD* universellen Primer sowie einer *pmoA2*-spezifischen Sonde (Anwendung nicht gezeigt). Die Optimierung der PCR-Protokolle, insbesondere der Anlagerungstemperaturen der einzelnen Systeme erfolgte anhand von aus Reinkulturen (*nif*-Primer) bzw. Klonen (PRO*pmoA2*) extrahierter DNA im Gradientencycler.

2.11.7 tRFcut

Im Rahmen der vorliegenden Arbeit wurde eine zusätzliche Funktionalität ("Add-In", "Tool") für das Programm-Paket ARB (Ludwig et al., 2004) entwickelt und implementiert: „tRFcut“. ARB ist eine Graphikorientierte modulare Sequenz-Daten-Verwaltung (vgl. 2.11.2) und wurde von Ludwig und Mitarbeitern in den letzten ca. 10 Jahren kontinuierlich weiterentwickelt. ARB wurde in der Programmiersprache Perl implementiert.

Um eine reibungslose Verteilung des Programms tRFcut an potentielle Nutzer zu ermöglichen, und um mögliche Beeinträchtigungen der bereits vorhandenen ARB-Funktionalität zu vermeiden, wurde tRFcut nicht direkt in den Source-Code von ARB integriert. Stattdessen wurde das Programm in Form zweier Shell-Skripte sowie einiger zusätzlicher Dateien realisiert. Die Skripte werden vom Nutzer mit individuell zu spezifizierenden Parametern direkt aus ARB heraus gestartet und erzeugen (bzw. löschen) nutzerspezifische individuelle tRF-Befehle. Diese automatisch generierten Befehle sind die eigentlichen Anweisungen für die *in silico* tRFLP. Sie werden in der rudimentären „arb-command-interpreter“ (ACI)-"Sprache" formuliert und in einer von ARB verwalteten Datei gespeichert. Jeder Befehl entspricht dabei genau einer vom Nutzer spezifizierten Kombination aus Primern und Endonuklease. Der Anwender kann die Befehle anschließend innerhalb der gewohnten ARB-Umgebung ausführen lassen, um *in silico* tRFLP-Analysen für beliebige Nukleotid-Sequenzen durchzuführen. Die Ergebnisse werden in entsprechenden Datenfeldern abgelegt und können für alle Datenbank-Anwendungen genutzt werden. Eine ausführliche Erläuterung von tRFcut ist unter 2.4.1 (Ergebnisse) sowie 4.4 (Diskussion) dargestellt.

tRFcut wurde mit allen verfügbaren ARB-Versionen getestet und läuft stabil in Kombination mit Versionen ab Release-Datum 1999. tRFcut wurde auf verschiedenen UNIX-basierten Betriebssystemen (Linux, Solaris) getestet. Lediglich die bei diesen Systemen standard-mäßig installierte "bash" (Bourne-again Shell) wird benötigt.

2.11.9 ORF-Vorhersage und Gen-Annotation

Die Primär-Annotation größerer genomischer Fragmente wurde automatisiert am MPI für molekulare Genetik, Berlin durchgeführt. Die Vorhersage offener Leserahmen ("ORF-Prediction") erfolgte mittels des Programms ORPHEUS (<http://pedant.gsf.de/orpheus/>). ORPHEUS kombiniert Datenbank-Vergleiche mit statistischen Analysen kodierender Sequenzabschnitte sowie möglicher Promoter-Bereiche um glaubwürdige Unterscheidungen tatsächlich transkribierter Gene von falsch-positiven ORFs zu ermöglichen. Vorhergesagte Gene bzw davon abgeleitete Aminosäuresequenzen wurden anschließend mittels der HTGA-Software (Rabus et al., 2002) gegen eine nicht-redundante DNA-Datenbank (NCBI), eine nichtredundante Protein-Datenbank (SWISSPROT, TREMBL und PIR) sowie die

Einträge der Interpro-Datenbank des EBI (<http://www.ebi.ac.uk/interpro/>) verglichen. HTGA ermöglicht somit nicht nur die Identifikation von zum untersuchten Genom-Abschnitt homologen Sequenzen, sondern darüberhinaus auch die Zuordnung der vorhergesagten Proteine zu Proteinfamilien oder COGs ("Clusters of Orthologous Groups") sowie das Erkennen funktionell hochkonservierter Protein-Domänen bzw. Sequenz-Muster. Mittels des Programms Artemis, (Releases 5 und 6) (<http://www.sanger.ac.uk/Software/Artemis/>) erfolgte anschließend am MPI für terrestrische Mikrobiologie, Marburg die Visualisierung der vorhergesagten Gene, die manuelle Feinannotation sowie der Vergleich der partiellen Genome mit öffentlich zugänglichen vollständigen Genom-Sequenzen hinsichtlich übereinstimmender konservierter genomischer Bereiche.

2.11.10 Phylogenetische Zuordnung anhand genomischer Signaturen

Die XanaMap Software (www.xanagene.com) erlaubt die phylogenetische Zuordnung kurzer genomischer Fragmente auf Basis Genom-intrinsischer Signaturen. Dazu wird eine statistische Analyse der Di-, Tri- und Tetranukleotide der zu analysierenden Fragmente vorgenommen und gegen die Signaturen der öffentlich zugänglichen komplett-sequenzierten Genome verrechnet. In der Literatur sind verschiedene Ansätze beschrieben, Gen-unabhängige genomische Charakteristiken vergleichend zu analysieren (Deschavanne et al., 1999; Edwards et al., 2002; Karlin et al., 1994; Karlin et al., 1992; Karlin et al., 1997; Knight et al., 2001; Oliver et al., 1993; Sandberg et al., 2001, Sandberg et al., 2003; Teeling et al., 2004). Eine besondere Bedeutung kommt dabei der anschließenden Gruppierung (Clustering) von Fragmenten mit zueinander ähnlichen Signaturen zu. XanaMap verwendet den Neuronales-Netzwerk-Algorithmus SOM (self-organizing map, Übersicht in: Kohonen, 1990) für das Clustering und die Visualisierung der Daten des hoch-dimensionalen Datenraums. Der ursprüngliche SOM-Algorithmus wurde dabei von Abe und Mitarbeitern (2002, 2003) an die Bearbeitung von genomischen Daten angepasst.

Vereinfacht dargestellt arbeitet der SOM-Algorithmus wie folgt. In einem ersten Schritt (Modell-Training) wird eine 2-dimensionale Matrix mit den genomischen Signaturen öffentlich verfügbarer Genome unüberwacht trainiert. Dabei

entspricht jeder Punkt der Matrix einem konkreten Datum im n-dimensionalen Vektorraum. Nach einer Initialisierung der Matrix, die bei XanaMap PCA- (Principal Component Analysis) basiert erfolgt, werden einzelne genomische Fragmente in randomisierter Reihenfolge mit der Matrix verrechnet. Dabei wird der zu dem zu verrechnenden Fragment ähnlichste Bereich der Matrix nach definierten Formeln an die tatsächliche Signatur des Fragments angepasst (der Grad der Anpassung ist dabei proportional zum Grad der Ähnlichkeit). Die anderen Bereiche bleiben unverändert. Mit fortwährendem Lernen passt sich die Matrix (bzw. die den einzelnen Punkten der Matrix zugeordneten Daten des hoch-dimensionalen Vektorraums) an die verrechneten Signaturen an. Die Matrix repräsentiert nun die in den Genomen tatsächlich vorkommenden Bereiche des Vektorraums in der Art, dass nahe bei einander liegende Punkte der Matrix hohe gemittelte Ähnlichkeitswerte aufweisen. In einem zweiten Schritt werden die Originaldaten der Matrix zugeordnet. Dabei werden aufgrund der Struktur der Matrix einander ähnliche Fragmente nah bei einander positioniert, unähnliche Fragmente entsprechend entfernt. Anschließend ist es möglich, solche Bereiche zu definieren, denen ausschließlich (bzw. überwiegend) genomische Fragmente eines einzigen Taxons zugeordnet wurden. Gegen diese Matrix können anschließend die zu analysierenden, beispielsweise aus Metagenom-Ansätzen stammenden, genomischen Fragmente verrechnet werden.

Für die SOM-Analyse wurden die in dieser Arbeit sequenzierten Genom-Abschnitte methanotropher Bakterien in 5 kb-Fragmente zerlegt (Sliding-Window-Methode: Fragment 1: Basen 1-5000, Fragment 2: Basen 1001-6000 usw.) und die so entstandenen Sub-Fragmente gegen die in der Software XanaMap vorhandenen Daten (5 kb-Fragmente) aller in GenBank vorhandenen vollständig sequenzierten Genome verrechnet. In die Analyse wurden auch die öffentlich verfügbaren, nicht redundanten Sequenzen von *Methylosinus* spp. bzw. *Methylocystis* spp. aufgenommen. Die Daten wurden vom NCBI bezogen, Sequenzen mehrfach sequenzierter homologer Abschnitte (beispielsweise partielle *pmoA*-Gene) auf eine repräsentative Sequenz reduziert und die Sequenzen einzeln gegen die XanaMap-Daten verrechnet. Darüberhinaus wurden die Sequenzen konkateniert, die entstandenen Konkatemere in 5 kb-Sub-Fragmente unterteilt und analog verrechnet.

2.11.11 Weitere verwendete Programme

Neben den bereits ausführlich erläuterten bioinformatischen Programmen, Online-Werkzeugen und Algorithmen wurden im Laufe dieser Arbeit eine Vielzahl weiterer Programme und Hilfsmittel genutzt. Eine Darstellung der zugrundeliegenden Algorithmen würde jedoch den gegebenen Rahmen deutlich übersteigen. Deshalb werden nachfolgend die für die durchgeführten Arbeiten wichtigsten Programme lediglich zusammenfassend vorgestellt.

Die Programme EditSeq und Seqman der kommerziellen Software-Suite DNA-Star[®] wurden für die manuelle Basenannotation sowie das Verbinden ("einrastern") einzelner Sequenzen zu größeren "Contigs" im Rahmen kleinerer, im Hause durchgeführter Sequenzierprojekte, beispielsweise der Sequenzanalyse des konventionellen pMMO-Operons von *Methylocystis* sp. Stamm SC2 verwendet.

Globale und lokale paarweise Alignments wurden mit den Programmen Stretcher und Matcher des Instituts Pasteur, Paris (<http://bioweb.pasteur.fr/seqanal/interfaces/>) durchgeführt. Hierbei handelt es sich um speicherplatzoptimierte Abwandlungen klassischer dynamischer Alignment-Algorithmen (Needleman und Wunsch, 1970; Smith und Waterman, 1981). Während globale Alignment-Algorithmen stets beide zu vergleichende Sequenzen über ihre gesamte Länge alignen, suchen lokale Alignment-Algorithmen lediglich den höchstkonservierten Bereich beider Sequenzen.

Die Identifikation homologer genomischer Abschnitte ("Gencluster") erfolgte mittels der kommerziellen Software ERGO sowie der Datenbanken "Clusters of Orthologous Groups of Proteins" des NCBI (<http://www.ncbi.nlm.nih.gov/COG/>) sowie der Datenbank KEGG-Orthology (<http://www.genome.jp/kegg/kegg2.html>).

Datenbank-Suchen wurden entweder mittels der diversen Blast-Programme des NCBI (Altschul et al., 1990 ; <http://www.ncbi.nlm.nih.gov/BLAST/>) oder mittels des FastA-Algorithmus am EBI/EMBL (Pearson, 1990; <http://www.ebi.ac.uk/fasta33/>) durchgeführt. Stichwort-basierte Datenbankabfragen wurden hauptsächlich mittels Entrez (NCBI), SRS (EBI) bzw der entsprechenden Werkzeuge der Protein- und Sekundärdatenbanken durchgeführt.

"Codon Usage"-Analysen wurden über das Web-Interface der Codon Usage Database (<http://www.kazusa.or.jp/codon/>) sowie des Programms Graphical Codon Usage Analyser (<http://gcua.schoedl.de/>) bzw. mittels der in anderweitig beschriebenen Programmen implementierten Algorithmen (Artemis, EditSeq) durchgeführt.

2.11.12 Computergestützte Vorhersage von Promotoren und Terminatoren

Die für *Methylocystis* sp. Stamm SC2 experimentell nachgewiesenen Transkriptions-Startpunkte von *pmoCAB1* und *pmoCAB2* ermöglichten den ungefähren Nachweis der Promoterregion. Eine Kontrolle sowie die finale basengenaue Bestimmung erfolgte mit Hilfe des Programms BPROM (<http://www.softberry.com/berry.phtml>). Faktorunabhängige Terminatoren wurden mittels des Programms TERMINATOR (Brendel und Trifonov, 1984) des "Genetics Computer Group (GCG) Software Package" vorhergesagt.

2.12 Oligonukleotide und Primer

In der folgenden Tabelle sind alle in dieser Arbeit mehrfach verwendeten Primer aufgeführt. Primer, die nur für eine einzige Anwendung formuliert wurden (beispielsweise über 100 verschiedene Sequenzier-Primer für das Primer-Walking) oder deren Anwendung keinen Erfolg erzielte, werden nicht aufgeführt. Der hier verwendete, von der IUBMB (International Union of Biochemistry and Molecular Biology) genormte 1-Buchstaben-DNA-Kode lautet wie folgt: M: A/C; R: A/G; W: A/T; S: G/C; Y: C/T; K: G/T; V: A/G/C; H: A/C/T; D: A/G/T; B: G/C/T; N: A/T/G/C. Alle in dieser Arbeit verwendeten Oligonukleotide wurden von der Fa. MWG Biotech AG (Ebersberg) in HPSF (High Purity Salt Free) – Qualität bezogen.

Tabelle 2.5: Verwendete Oligonukleotid-Primer

Primer	Position ⁴	Sequence (5'→3')	Direction
<u><i>pmoA</i>-Primer</u>			
A189f ¹	1486-1504 [*]	GGN GAC TGG GAC TTC TGG	Forward
PmoA206f ²	1486-1521 [#]	GGNGACTGGGACTTCTGGATCGACTTCAAGGATCG	Forward
682b ²	2026-2009 [*]	GAA SGC NGA GAA GAA SGC	Reverse
PmoA703b ¹	2026-1992 [#]	GAASGCNGAGAAGAASGCGGCGACCGGAACGACGT	Reverse
<u><i>nifH</i> / <i>nifD</i>-Primer</u>			
nifH-F1		TAYGGNAARGGNGGNATYGGNAARTC	Forward
nifH-r		ADNGCCATCATYCTNCC	Reverse
nifHD-f	436-461 des <i>B. japonicum</i> USDA 110 <i>nifH</i> -gens	CAGGAAATCTACATCGTCATGTC	Forward
nifD-f	208-227 des <i>B. japonicum</i> USDA 110 <i>nifD</i> -gens	GYGGYTGCCTAYGCCGG	Forward
nifD-r		TCCCANGARTGCATCTGRCGGA	Reverse
<u>RACE-Primers</u>			
NovC316r	511-492 [#]	CAG CCG ATG AGC GTC AGG AT	Reverse
NovC100r	293-273 [#]	GCC GAC GTA GAA GAC CAG CAT	Reverse
NovC40r	236-216 [#]	GAT CGT ATC GGT CTC GCC GGC	Reverse
ConvC100r	445-425 [*]	TCG TAA ATG CGA ACG ACC AGA	Reverse
ConvC40r	372-352 [*]	CAG CTT CCG TGC CAG CAG CAG	Reverse
ConvC-60r	269-289 [*]	TGC GCG TGA CCG GTA TCG CTC	Reverse
<u>RACE-Kit-Primer</u>			
Oligod(T) Anchored ³	-38 - -1 ^{*#}	GAC CAC GCG TAT CGA TGT CGA C(T) ₁₆ V	Forward
Anchor primer ³	-38 - -17 ^{*#}	GAC CAC GCG TAT CGA TGT CGA C	Forward
<u>Sequenzier-Primer:</u>			
M13F		GTT TTC CCA GTC ACG AC	-
M13R		CAG GAA ACA GCT ATG AC	-
<u>Formid-spezifische-Primer:</u>			
pCC1FOS-FP ⁵		GGA TGT GCT GCA AGG CGA TTA AGT TGG	-
pCC1FOS-RP ⁵		CTC GTA TGT TGT GTG GAA TTG TGA GC	-

¹ Holmes et al. (1996)

² Tchawa Yimga et al. (2003)

³ Roche Diagnostics (Mannheim)

⁴ (*)^(#) Primer Positionen sind in Relation zum Transkriptionsstartpunkt von *pmoCAB1*(*) bzw *pmoCAB2* (#) angegeben.

⁵ Epicentre (Madison, USA)

3. Ergebnisse

3.1 Untersuchungen der für pMMO-2 kodierenden Gene

3.1.1 Untersuchungen zu *pmoA2*, einer neuartigen Genkopie der *pmoA*

Übersicht :

Um die Verbreitung von *pmoA2* bei methanotrophen Bakterien zu untersuchen, wurden 32 Typ I und Typ II MB mittels tRFLP-Analyse sowie mittels eines neu entwickelten *pmoA2*-spezifischen PCR-Assays untersucht. Die Analyse ergab eine weite Verbreitung von *pmoA2* innerhalb der Typ II MB. Im Gegensatz dazu konnte das Gen in keinem der fünf untersuchten Typ I MB nachgewiesen werden. Interessanterweise fehlt das *pmoA2*-Gen auch in Genomen mancher Typ II MB, obwohl es für phylogenetisch nah verwandte Organismen nachgewiesen werden konnte. Da die mangelnde Nachweisbarkeit auf schlecht bindende Primer zurückführbar sein könnte, wurde mittels Southern-Hybridisierungs-Experimenten ausgeschlossen. In seinem Verbreitung-Muster erinnert *pmoA2* deshalb an die sMMO.

In phylogenetischen Verrechnungen gegen Sequenzen konventioneller *pmoA*-Gene bilden die partiellen *pmoA2*-Sequenzen eine gemeinsam verzweigende Entwicklungslinie. Diese Line gruppiert gemeinsam mit den konventionellen *pmoA*-Sequenzen der Typ II MB. Sowohl Analysen der synonymen vs. nicht-synonymen akzeptierten Mutationen wie auch der RT-PCR-basierte Nachweis *pmoA2*-spezifischer cDNA resultierten in deutlichen Hinweisen auf eine funktionelle Relevanz von PmoA2.

Anmerkung :

Folgende Teile der nachfolgend reproduzierten Publikation wurden im Rahmen dieser Promotion bearbeitet: anteilige Untersuchung der genannten MB auf Vorhandensein von *pmoA2*, Erhebung der genannten Sequenz-Daten sowie phylogenetische Analyse der erzielten Sequenzen.

Wide Distribution of a Novel *pmoA*-Like Gene Copy among Type II Methanotrophs, and Its Expression in *Methylocystis* Strain SC2

Merlin Tchawa Yimga, Peter F. Dunfield, Peter Ricke, Jürgen Heyer, and Werner Liesack*

Max-Planck-Institut für terrestrische Mikrobiologie, D-35043 Marburg, Germany

Received 7 March 2003/Accepted 30 June 2003

Experiments were conducted to determine if a novel *pmoA*-like gene (*pmoA2*) recently discovered in the methane-oxidizing bacterium *Methylocystis* strain SC2 (P. F. Dunfield, M. Tchawa Yimga, S. D. Dedysh, U. Berger, W. Liesack, and J. Heyer, FEMS Microbiol. Ecol. 41:17–26, 2002) is present in other methane-oxidizing bacteria (MOB), and if it is expressed. A newly developed primer combination (*pmoA206f*-*pmoA703b*) allowed a differential detection of *pmoA1* and *pmoA2*. By using this primer combination, we identified *pmoA2* in a wide range of type II MOB of the *Methylosinus*-*Methylocystis* group. However, screening by PCR and by Southern hybridization using a newly developed *pmoA2*-specific oligonucleotide probe also showed that closely related type II MOB, exhibiting 16S rRNA gene sequence identities of higher than 97%, may or may not harbor *pmoA2*. No *pmoA2* was detected in five type I MOB tested: *Methylococcus capsulatus* strain Bath, *Methylocaldum* strain E10A, *Methylobacter luteus*, *Methylobacterium album*, and *Methylomonas* strain D1a. In comparative sequence analyses, all *pmoA2*-like sequences formed a coherent cluster clearly distinct from *pmoA1* sequences of type I and type II MOB, and from *amoA* sequences of the *Nitrosomonas*-*Nitrospira* group. Phylogenetic analysis using the paml model suggested that *pmoA2* is subject to strong purifying selection and therefore has an important cellular function. We probed total RNA extracts of *Methylocystis* strain SC2 for gene expression of *pmoA*. A strong signal was observed for *pmoA1* in Northern hybridization, while the results obtained for *pmoA2* were ambiguous. However, reverse transcription-PCR confirmed that *pmoA2* was expressed, albeit at lower level than *pmoA1*. This provided experimental evidence that the gene product of *pmoA2* may be a functionally active enzyme.

Methane-oxidizing bacteria (MOB) are able to utilize methane (CH₄) as their sole source of carbon and energy for growth (13). Their ability to oxidize CH₄ released at the interface of methanogenic environments and to act as sink for atmospheric CH₄ makes these organisms key players in balancing the global CH₄ budget and mitigating global warming due to CH₄ (4, 28). Phylogenies based on 16S rRNA genes show that MOB form distinct lineages in the *Gammaproteobacteria* (type I MOB) and *Alphaproteobacteria* (type II MOB) (2, 5, 6, 13, 16).

The first step in CH₄ oxidation, the conversion of methane to methanol, is carried out by a methane monooxygenase (MMO). This enzyme exists in two forms, a particulate, membrane-associated form (pMMO) and a soluble form (sMMO). The two forms of enzyme differ in their structures, kinetic properties, and ranges of substrates they utilize (26). Only a restricted number of MOB species harbor sMMO, while almost all MOB possess pMMO. The only MOB lacking pMMO is *Methylocella palustris* (5).

Cloning and sequence analysis of genes encoding pMMO revealed three consecutive open reading frames (*pmoC*, *pmoA*, and *pmoB*) in both type I (30, 31) and type II MOB (11). The *pmoA* gene, which encodes the 27-kDa subunit (PmoA) of pMMO, has been shown to be evolutionarily highly conserved among methanotrophs (17). The type I MOB *Methylococcus capsulatus* strain Bath and *Methylobacterium album* strain

BG8 (30, 31), as well as the type II species *Methylosinus trichosporium* strain OB3b and *Methylocystis* sp. strain M (11) contain duplicate *pmoA* gene copies. The sequences of these duplicate gene copies have been shown to be nearly identical.

However, we recently demonstrated that the type II MOB *Methylocystis* strain SC2 contains two very different *pmoA*-like genes (9). The first gene (*pmoA1* or conventional *pmoA*) exhibited very high sequence homology to *pmoA* genes of other type II MOB (even identical amino acid sequence to PmoA of some other *Methylocystis* strains). The second gene (*pmoA2* or novel *pmoA*) possessed only 73% identity with *pmoA1* at the nucleotide level and 68.5% identity at the amino acid level. *PmoA2* of *Methylocystis* strain SC2 was closely related to the deduced amino acid sequence of a *pmoA*-like gene retrieved in a previous study by cultivation-independent methods from rice field soil (86.3% identity) (19), indicating that the presence of multiple, diverse *pmoA* gene copies might not be unique to strain SC2. This was also demonstrated by the detection of a novel *pmoA2*-like copy in *M. trichosporium* strain KS21 (9).

Our study aimed to assess whether the presence of multiple, diverse *pmoA* gene copies is a common genotypic trait among methanotrophs or a phenomenon restricted to a few methanotroph strains. We developed PCR primers and hybridization probes for differential detection of *pmoA1* and *pmoA2* in single MOB genospecies. PCR-based screening by terminal restriction fragment length polymorphism (T-RFLP) analysis, Southern hybridization and comparative sequence analysis revealed that the presence of multiple, diverse *pmoA* genes in single genospecies is common among type II MOB, but not type I MOB. We also show that, contrary to previous results

* Corresponding author. Mailing address: Max-Planck-Institut für terrestrische Mikrobiologie, Karl-von-Frisch-Str., D-35043 Marburg, Germany. Phone: 49 (6421) 178 720. Fax: 49 (6421) 178 809. E-mail: liesack@staff.uni-marburg.de.

TABLE 1. List of methanotrophic strains screened for the presence of *pmoA2*

Strain designation	Taxonomic assignment	T-RFs at temp ^a :		Detection of <i>pmoA2</i>	Source or reference
		60°C	66°C		
SM16	<i>Methylocystis</i> sp.	245, 438	438	+	16
LR1	<i>Methylocystis</i> sp.	245, 438	438	+	8
B2/7	<i>Methylocystis</i> sp.	245, 438	438	+	16
62/12	<i>Methylocystis</i> sp.	245, 438	438	+	16
KS9	<i>Methylocystis</i> sp.	245, 350	350	+	16
O14c	<i>Methylocystis</i> sp.	245	245	—	16
IMET 10484	<i>Methylocystis</i> sp.	245, 350	350	+	16
21/1	<i>Methylocystis</i> sp.	245	245	—	16
IMET 10499	<i>Methylocystis</i> sp.	245, 350	350	+	16
F10V2a	<i>Methylocystis</i> sp.	115, 245	115	+	16
KS12	<i>Methylocystis</i> sp.	209	209	—	16
IMET 10486	<i>Methylocystis</i> sp.	245, 438	438	+	16
SC2	<i>Methylocystis</i> sp.	245, 438	438	+	16
IMET 10491 ^T	<i>Methylocystis echinoides</i>	245, 438	438	+	16
Pi5/4	<i>Methylocystis</i> sp.	245	245	—	16
81	<i>Methylocystis parvus</i>	245, 350	350	+	16
SC8	<i>Methylosinus sporium</i>	245, 350	350	+	16
SK13	<i>Methylosinus sporium</i>	245, 350	350	+	16
20/3	<i>Methylosinus sporium</i>	245, 159	159	+	16
H1b	<i>Methylosinus sporium</i>	245, 350	350	+	16
SM6	<i>Methylosinus trichosporium</i>	245, 350	350	+	16
H5	<i>Methylosinus trichosporium</i>	245	ND ^b	—	16
M23	<i>Methylosinus trichosporium</i>	245	245	—	16
39/3	<i>Methylosinus trichosporium</i>	245	245	—	16
NCIMB 11131 ^T , OB3b	<i>Methylosinus trichosporium</i>	245	ND	—	NCIMB ^c
SC10	<i>Methylosinus trichosporium</i>	130, 350	350	+	16
KS21	<i>Methylosinus trichosporium</i>	245, 279	279	+	16
NCIMB 11914 ^T	<i>Methylobacter luteus</i>	ND	ND	—	NCIMB
NCIMB 11123 ^T , BG8	<i>Methylobacterium album</i>	ND	ND	—	NCIMB
NCIMB 11853, Bath	<i>Methylococcus capsulatus</i>	ND	ND	—	NCIMB
D1a ^d	<i>Methylobacter</i> sp.	ND	ND	—	This study
E10a ^d	<i>Methylocaldum</i> sp.	ND	ND	—	This study

^a T-RFs observed in differential T-RFLP analysis. PCR at an annealing temperature of 60°C enabled simultaneous detection of both *pmoA1* and *pmoA2*, while PCR at 66°C led to the specific PCR amplification of *pmoA2* (if present).

^b ND, not detected.

^c NCIMB, National Collections of Industrial and Marine Bacteria Ltd.

^d *Methylobacter* strain D1a and *Methylocaldum* strain E10a are novel isolates of type I MOB. Their taxonomic assignment is based on comparative 16S rRNA gene sequence analysis. Further details of the two strains will be published elsewhere.

(9), *pmoA2* of *Methylocystis* strain SC2 is expressed under standard growth conditions, although the level of expression is clearly lower than that of *pmoA1*.

MATERIALS AND METHODS

MOB strains. The 32 strains used in this study are listed in Table 1. The conditions used for growth of the MOB were adapted from Heyer et al. (15, 16) or Leadbetter and Forster (22) and depended upon whether the cultures were intended for screening of *pmoA2* by PCR amplification or Southern blot analysis. Cultures to be used for PCR-based screening were grown, as described by Heyer et al. (16), on agar plates of mineral salts medium incubated at 30°C in closed glass chambers containing a gas mixture of 20% (vol/vol) CH₄, 5% CO₂, and 75% air. For Southern or Northern hybridization experiments, MOB were grown in liquid culture of medium 10 for *M. trichosporium* strain OB3b and *Methylocystis* strain SC2 or in medium 10 containing NaNO₃ instead of NH₄Cl as the nitrogen source for *M. capsulatus* strain Bath and *Methylobacter* strain D1a. For growth of *Methylobacter* strain D1a, a supplement of vitamin solution was added to final concentrations of 1 µg/liter for vitamin B₆; 0.2 µg/liter for biotin and folic acid; 0.5 µg/liter for vitamin B₁, vitamin B₂, and DL-Ca-pantothenate; and 0.01 µg/liter for vitamin B₁₂. Liquid cultures were grown for 3 to 5 days at 30°C under a headspace of 20% (vol/vol) CH₄, 5% CO₂, and 75% air. Cells were pelleted at 4,000 × g for 20 min at 4°C and washed once with TE buffer (10 mM Tris, 1 mM EDTA [pH 8.0]). The cell biomass was immediately used for extraction of DNA.

DNA extraction. DNA extraction for PCR-based studies was performed by a procedure based on mechanical agitation in a FastPrep FP120 cell disrupter (Savant, Holbrook, N.Y.) of 2-ml screw-cap reaction vessels filled with a mixture

of culture, 0.1-mm-diameter silica-zirconium beads and phosphate buffer (pH 8.0) containing 2% sodium dodecyl sulfate (SDS) (14).

Cells grown in liquid culture (~1 g of fresh biomass) were suspended in 10 ml of TE buffer to which 25 mg of lysozyme and 20 µl of proteinase K (25 mg/ml) were added. The suspension was placed in a 37°C water bath for 2 h, after which a 10% (wt/vol) solution of SDS was added to a final concentration of 1%, followed by a 1.5-h incubation at 37°C. After centrifugation for 15 min at 4°C, the supernatant was collected, and 1 ml of 5 M potassium acetate (pH 7.5) was added. The suspension was centrifuged for 15 min at 4°C, and the resulting supernatant was transferred to a new vessel. Total nucleic acids were extracted from the supernatant twice with 1 volume of chloroform-isoamyl alcohol (24:1 [vol/vol]) and then precipitated from the aqueous phase with 1 volume of isopropanol. The nucleic acids were resuspended in 5 ml of TE buffer.

For removal of coextracted RNA, 30 µl of 100-mg/ml RNase A and 20 µl of 100,000-U/ml RNase T₁ were added, followed by incubation for 2 h at 37°C. The RNase treatment was stopped by extraction three times with 1 volume of chloroform-isoamyl alcohol (24:1 [vol/vol]). The DNA was precipitated from the aqueous phase with 1 volume of isopropanol and then resuspended in 100 µl of TE buffer. The amount of DNA extracted was estimated by electrophoresis of aliquots on a 1% agarose gel and comparison to a *Pst*I digest of phage λ DNA.

Extraction of total RNA from *Methylocystis* strain SC2. Total nucleic acids were extracted as described above for strains grown in liquid culture, except that solutions for extracting RNA were prepared with diethyl pyrocarbonate (DEPC)-treated deionized water. For the removal of coextracted DNA, 1 volume of TMC buffer (10 mM Tris-HCl [pH 7.5], 5 mM MgCl₂, 0.1 mM CsCl₂) (7) and 20 µl (5 U) of RNase-free DNase (Promega, Madison, Wis.) were added, followed by incubation for 2 h at 37°C. The reaction was stopped by extraction

TABLE 2. Oligonucleotide primers and probes used in this study

Name ^a	Target	Strategy	Gene positions ^b	Sequence (5'→3') ^c	Reference
A189f	<i>pmoA</i>	PCR, RT-PCR	172–189	GGNGACTGGGACTTCTGG	17
A682b	<i>pmoA</i>	PCR, RT-PCR	702–685	GAAGCGNGAGAAGAASGC	17
pmoA206f	<i>pmoA1</i> and <i>pmoA2</i> ^d	PCR, RT-PCR	172–206	GGNGACTGGGACTTCTGGATCGACTTCAAGGATCG	This study
pmoA703b	<i>pmoA1</i> and <i>pmoA2</i> ^d	PCR	702–668	GAAGCGNGAGAAGAASGCAGCGGACCGGACGACGT	This study
pmoA636b	<i>pmoA2</i>	S/N blot, ^e RT-PCR	635–610	ATCATGCGGATGTATTTCMGGSGTGCC	This study
A593b	<i>pmoA1</i> (strain SC2)	S/N blot, RT-PCR	615–593	CATCGACGTGCGGACGAAGTGA	9
Eub9f	16S rRNA gene	PCR	9–27	GAGTTTGATCMTGGCTCAG	21
Eub1492b	16S rRNA gene	PCR	1512–1492	ACGGYTACCTTGTACGACTT	34

^a All (including labeled) oligonucleotides were purchased from Metabion (Martinsried, Germany).

^b Numbering for *pmoA* refers to the *pmo* gene sequence of *Methylococcus capsulatus* strain Bath (30). Numbering for 16S rDNA refers to the 16S rRNA gene sequence of *E. coli* (3).

^c N = A, T, C, or G; M = C or A; R = A or G; Y = C or T; and S = G or C.

^d The use of the primer combination pmoA206f and pmoA703b in PCR at an annealing temperature of 60°C enabled simultaneous detection of both *pmoA1* and *pmoA2*, while use at 66°C led to the specific PCR amplification of *pmoA2* (if present).

^e S/N blot, Southern or Northern blot.

three times with 1 volume of chloroform-isoamyl alcohol (24:1 [vol/vol]). The total RNA was precipitated from the aqueous phase with 1 volume of isopropanol and resuspended in 100 µl of DEPC-pretreated TE buffer. The integrity of the 16S and 23S rRNA fragments was checked by electrophoresis on a 1.2% agarose gel and comparison to an rRNA standard from *Escherichia coli* (Roche Diagnostics).

Differential detection of *pmoA1* and *pmoA2*. The *pmoA*-based diversity present in a single strain was assessed by determining the number and the size of terminal restriction fragments (T-RFs) observed in restriction digests of *pmoA* genes PCR-amplified with the primer combination pmoA206f-pmoA703b at two different annealing temperatures (Table 2). For T-RFLP analysis, the 5' primer (pmoA206f) was labeled with the dye 5-carboxyfluorescein. The reaction mixture contained 1 µl of template DNA, 7.5 µl of 10× reaction buffer, 1.5 mM Mg²⁺, 150 µM (each) deoxynucleoside triphosphate (dNTP), 0.125 µM (each) primer, and 2.5 U of *Taq* DNA polymerase (Promega, Mannheim, Germany). The thermal PCR profile was as follows: initial denaturation for 3 min at 94°C followed by 32 cycles consisting of denaturation at 94°C for 60s, primer annealing at 60 or 66°C for 60 s, and elongation at 72°C for 60s. The final extension step was extended to 7 min. Aliquots of the amplicons (10 µl) were checked by electrophoresis on a 1% agarose gel.

The digestion of purified amplicons (100 ng) and separation of the restriction digests on an ABI 373A automated sequencer (Perkin-Elmer Applied Biosystems, Weiterstadt, Germany) were carried out as described previously (24).

For direct sequencing of amplicons obtained at an annealing temperature of 66°C, PCR was carried out with primer pmoA206f without fluorescence label. Amplicons were purified by using the Qiagen PCR purification kit (Qiagen, Hilden, Germany) according to the recommendations of the supplier. Both strands of the PCR products were analyzed by using the dye terminator sequencing chemistry of PE Applied Biosystems and either pmoA206f or pmoA703b as the sequencing primer.

Southern hybridization. Aliquots of genomic DNA (10 µg) were digested overnight at 37°C with 50 U of restriction enzyme *Pst*I, *Eco*RI, or *Xho*I according to the protocols recommended by the supplier (Promega, Madison, Wis.). The restricted DNA was precipitated with ethanol, dried in a vacuum desiccator (Savant), resuspended in 20 µl of TE buffer, and separated on a 0.8% agarose gel at 25 V for 12 h. Southern blotting was carried out according to the procedure described in the *DIG Application Manual for Filter Hybridization* (Roche Diagnostics GmbH, Mannheim, Germany, 2000).

The digoxigenin (DIG)-labeled oligonucleotide probe pmoA636b (Table 2) was used for specific detection of *pmoA2*. A *pmo* gene probe was used for simultaneous detection of both *pmoA1* and *pmoA2*. This probe was based on a mixture of DIG-labeled amplicons of the different strains tested. PCR was carried out with either primer combination A189f-A682b (16) or pmoA206f-pmoA703b and an annealing temperature of 60°C. For direct labeling during PCR, DIG-dUTP was mixed with dTTP in the proportional relationship of 1:3.

Both oligonucleotide probe pmoA636b and the *pmo* gene probe were applied under the same stringency conditions. Hybridizations were carried out overnight at 42°C with a standard hybridization solution (5× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate], 0.1% *N*-lauroylsarcosine, 0.02% SDS, 1% blocking solution). Membranes were washed in 0.5× SSC at 60°C for 30 min, and hybridization signals were detected on a Storm 860 PhosphorImager (Molecular Dynamics, Sunnyvale, Calif.) after reaction with the ECF substrate for Western blotting (Amersham, Piscataway, N.J.).

Northern hybridization. Approximately 60 µg of total RNA from *Methylocystis* strain SC2 was first separated on a 1% formaldehyde agarose gel at 60 V for 3 h in presence of 1× MOPS (morpholinepropanesulfonic acid) buffer (10× MOPS buffer is 200 mM MOPS, 50 mM sodium acetate, 20 mM EDTA [pH 7.0]). Northern blotting was carried out according to the procedure outlined in the *DIG Application Manual for Filter Hybridization* (Roche Molecular Diagnostics). Blotting was performed overnight on a Hybond-N-membrane (Amersham). The DNA was UV cross-linked using a UV Stratelinker 2400 (Stratagene, La Jolla, Calif.) at 266 nm for 3 min. The DIG-labeled oligonucleotide probes A593b and pmoA636b were used for specific detection of *pmoA1* and *pmoA2*, respectively (Table 2). Hybridizations were performed overnight at 40°C. Washing and visualization were carried out as described above for Southern hybridization.

RT-PCR. Copy DNA (cDNA) of mRNA from *pmoA* genes expressed in *Methylocystis* strain SC2 was synthesized with a Qiagen Omniscript kit according to the instructions of the manufacturer. The reverse transcription (RT) was carried out in a total volume of 20 µl at 37°C for 30 min. The reaction mixture contained 2 µg of total RNA from strain SC2, 0.5 mM (each) dNTP, reverse transcriptase buffer, 10 U of RNase inhibitor, 1.0 µM primer A682b (Table 2), and 4 U of Omniscript reverse transcriptase. Aliquots (1 µl) of the cDNA solution were used for subsequent PCR amplification. The primer combinations A189f-A593b and pmoA206f-pmoA636b were used to specifically amplify cDNA from either *pmoA1* or *pmoA2*. The reaction cocktail contained 1 µl of RT product, 10 µl of 10× reaction buffer, 1.5 mM Mg²⁺, 200 µM each deoxynucleoside triphosphate (dNTP), 0.25 µM each primer, and 2.5 U of *Taq* DNA polymerase (Promega). The thermal profile was as follows: initial denaturation for 3 min at 94°C followed by 32 cycles of denaturation at 94°C for 40 s, primer annealing at 62°C for 40 s, and elongation at 72°C for 45s. The final extension step was extended to 7 min. Amplification was performed in a total volume of 100 µl in 0.2-ml reaction vessels and a DNA thermal cycler (model 2400; PE Applied Biosystems). Aliquots of the amplicons (10 µl) were checked by electrophoresis on a 1% agarose gel. Amplicons of the expected size were purified and sequenced as described above.

Analysis of sequence data. Trees were constructed for partial *PmoA* sequences (151 deduced amino acid positions) by using TreePuzzle and the ARB program package (developed by O. Strunk and W. Ludwig of the Technical University of Munich [http://www.arb-home.de]). TreePuzzle is a quartet maximum-likelihood method for reconstructing tree topologies (32, 33). The amino acid substitution model described by Adachi and Hasegawa (1) was used to construct the trees.

A tree was constructed for the same set of data on the nucleic acid level by the maximum-likelihood method FastDNAm1 (10).

In order to gain insight into the evolution and expression of the *pmoA2* gene, a phylogenetic tree of selected sequences was analyzed with the codeml function of the paml program of Ziheng Yang (http://abacus.gene.ucl.ac.uk/software/paml.html#GetPAML). This program calculates the relative rate of nonsynonymous (amino acid changing) to synonymous (non-amino acid changing) nucleotide substitutions (d_N/d_S , or ω) in a phylogeny (35). An ω value of <1 indicates that purifying natural selection is acting to prevent changes in a protein, while an ω value of >1 indicates diversifying selection. A ω value = 1 indicates neutral selection and would be expected for a pseudogene.

Nucleotide accession numbers. The *pmoA2* gene sequences obtained in the course of this study from environmental samples and cultured MOB have been deposited in the EMBL, DDBJ, and GenBank databases under the accession no. AJ543418 to AJ543423 and AJ544093 to AJ544102, respectively. The *pmoA1*

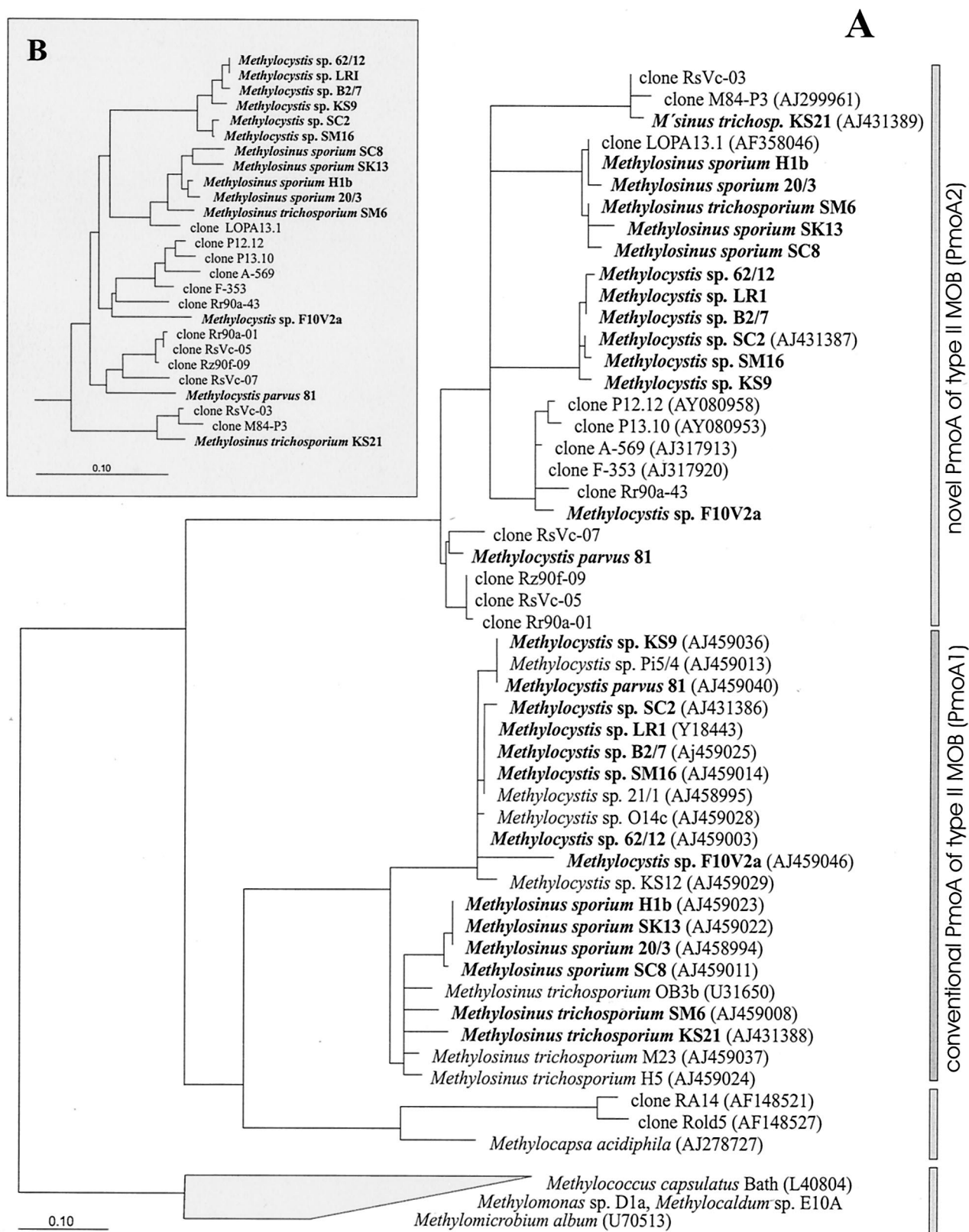


FIG. 1. Phylogenetic trees constructed for partial *pmoA* sequences and deduced amino acid sequences (PmoA1 and PmoA2). (A) PmoA tree showing the relationship of PmoA1 and PmoA2 of type II MOB of the *Methylosinus-Methylocystis* group to environmental clones as well as to *Methylocapsa acidiphila* and selected type I MOB. The environmental sequences were retrieved by cultivation-independent methods from the following habitats: forest soil (clones RA14, Rold5 [18], and P12.12 and P13.10 [27]), peat soil (clone LOPA13.1 [25]), natural and manipulated organic soils (clones A-569 and F-353), as well as bulk soil (clones RsVc-03, RsVc-05, and RsVc-07) and rice roots (clones M84-P3 [19], Rr90a-01,

sequences obtained for reference from *Methylocaldum* strain E10a and *Methylomonas* strain D1a have been deposited under accession no. AJ5441091 and AJ544092, respectively.

RESULTS

Differential detection of *pmoA1* and *pmoA2* by T-RFLP analysis. The first aim of the present study was the development of an appropriate method for detection of multiple, divergent *pmoA* genes present in a single organism. Using the *pmoA2*-like environmental clone sequence M84-P3 (19) and the *pmoA2* sequence of *Methylocystis* strain SC2 as the starting point, we designed several forward and reverse primers for specific retrieval of *pmoA2* from both environmental samples and cultured MOB. According to the *pmoA* database available, none of these primers perfectly matched *pmoA1* or *amoA*, suggesting that these primers should be specific for *pmoA2*. Environmental DNA from rice paddy soil is composed of genomic DNA from a complex community of phylogenetically diverse microorganisms, including various type I and II methanotrophs (19). Consequently, to assess the target specificity and the range of *pmoA2*-like sequence diversity detectable, we tested various primer combinations in first- and second-round PCRs using environmental DNA extracted from flooded rice microcosms in a previous study (19). This led to the retrieval of various *pmoA2*-like sequences from bulk soil and rice roots of flooded rice microcosms. Six representatives of these environmental clone sequences have been deposited in the EMBL, DDBJ and GenBank databases (see Materials and Methods) and are shown in Fig. 1. The best performance with regard to both target specificity and range of *pmoA2*-like sequence diversity detectable was exhibited by the primer combination pmoA206f-pmoA703b (Table 2).

Using *Methylocystis* strain SC2 as the model organism, this primer combination was therefore chosen to establish for cultured type II MOB a tool for differential detection of *pmoA1* and *pmoA2* by T-RFLP analysis. The T-RF size distribution of *pmoA1* and *pmoA2* was assessed by computer simulation. The T-RF sizes (5' or 3' termini) were predicted for as many as 10 different restriction endonucleases. For *MspI*, we identified two distinct 5'-T-RFs, which corresponded to *pmoA1* (245 bp) and *pmoA2* (438 bp), respectively. Based on the in silico analysis, we concluded that PCR with primer pair pmoA206f-pmoA703b followed by digestion with *MspI* was the simplest way to classify the two different *pmoA* sequence types present in *Methylocystis* strain SC2. To test our predictions experimentally, PCR amplification of *pmoA* was conducted for subsequent T-RFLP analysis. The effect of annealing temperature used in PCR on the *pmoA* sequence type detectable via T-RFLP analysis was investigated by performing the amplifications at various annealing temperatures. Two different T-RFs

with sizes of 245 and 438 bp were observed at an annealing temperature of 60°C, whereas only one distinct T-RF was observed at an annealing temperature of 66°C (Fig. 2). Thus, no discrepancy was observed between the results expected by in silico analyses and the empirical data obtained. However, to verify that only *pmoA2* was amplified at 66°C, another PCR was carried out with the annealing temperature of 66°C, and the resulting PCR product was sequenced. Comparative sequence analysis confirmed that the product was derived from *pmoA2*. This finding indicated that T-RFLP analysis in combination with comparative sequence analysis could be a rapid and reliable method for screening methanotrophic pure cultures for the presence of multiple, divergent *pmoA* gene copies.

Screening for *pmoA2* in single MOB strains by T-RFLP analysis. In addition to *Methylocystis* strain SC2 and *M. trichosporium* strain KS21, 30 pure cultures, including 25 type II MOB and 5 type I MOB, were screened for the presence of *pmoA2* by differential T-RFLP analysis. Prior to *pmoA*-based analysis, the identity of each strain was confirmed by comparative sequence analysis of its 16S rRNA gene (16).

Except for *M. trichosporium* strains OB3b and H5, PCR products were obtained for all type II MOB at both annealing temperatures (60 and 66°C). The five type I MOB tested failed to give a positive PCR with the primer combination pmoA206f-pmoA703b (Table 1). The T-RFLP patterns obtained for the type II MOB species revealed multiple, distinct T-RFs within single strains, suggesting that some degree of *pmoA* sequence diversity was present. For all of the strains that tested positive at an annealing temperature of 66°C, a single T-RF was observed. Depending upon the strain, these T-RFs were 115, 159, 209, 245, 279, 350, or 438 bp in size (Table 1). Representatives of each of these different T-RFs were further investigated by direct sequencing of the *pmoA* amplicon generated at an annealing temperature of 66°C. Translated sequences of PCR products were aligned and compared with a set of representative PmoA sequences. It was confirmed that the 209- and 245-bp T-RFs were derived from typical *pmoA1* sequences of type II MOB (*Methylosinus*-*Methylocystis* group) that could still be amplified with the presumed *pmoA2*-specific primer set at 66°C. However, T-RFs with sizes of 115, 159, 279, 350, and 438 bp were indicative of *pmoA2* (Table 1). The treeing analysis showed that all PmoA2 sequences obtained from type II MOB of the *Methylosinus*-*Methylocystis* group (Table 1) formed a coherent cluster clearly distinct from that of PmoA1 (Fig. 1). In addition, cloned sequences retrieved by cultivation-independent approaches from various environments, including forest soil, peat soil, natural and manipulated organic soils, and flooded rice microcosms grouped within the PmoA2 cluster.

The intercluster PmoA identity values between sequence

Rr90a-43, and Rr90f-09) of flooded rice microcosms. In the present study, the retrieval of environmental *pmoA2*-like sequences from the bulk soil and rice roots of flooded rice microcosms was based on total DNA extracted in a previous study (19). Type II MOB for which both PmoA1 and PmoA2 sequences could be detected are shown in boldface. GenBank accession numbers are given in parentheses after the species or clone names for sequences that were not obtained in this study. The tree was constructed by using TreePuzzle, a quartet maximum-likelihood method (32, 33). Lineages for which, based on 25,000 puzzling steps, the exact branching order could not be unambiguously determined are shown by multifurcation. The scale bar represents 0.1 change per amino acid position. (B) *pmoA* tree showing the *pmoA2* intracluster relationship at the nucleic acid level. The tree was constructed with the same data set as that on the PmoA level and by the maximum-likelihood method (FastDNAm1 [10]); however, the presentation of the tree is restricted to the *pmoA2* cluster. The distance bar represents 0.1 substitution per nucleotide position.

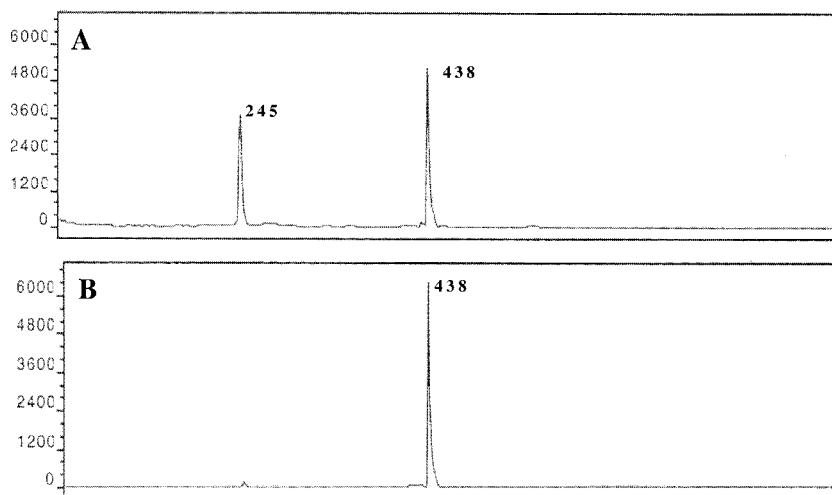


FIG. 2. *pmoA*-based T-RFLP profiles of *Methylocystis* strain SC2. PCRs were carried out with primers *pmoA*206f and *pmoA*703b at two different annealing temperatures: 60°C (A) and 66°C (B). On the y axis, the intensities of fragments are given in arbitrary units. Numbers indicate the lengths of T-RFs corresponding to *pmoA*1 (245 bp) and *pmoA*2 (438 bp).

types of PmoA1 and PmoA2 ranged from 64 to 74.5%. The lowest *pmoA*2 and PmoA2 intracluster identity values were 18.2 and 21.5%, respectively. For comparison, the corresponding values for *pmoA*1 and PmoA1 were 14.7 and 12.4%, respectively. The PmoA2 cluster was divided into five sublineages. These sublineages were characterized by *Methylocystis parvus* strain 81, *Methylocystis* strain F10V2a, *M. trichosporium* strain KS21, or by a set of either *Methylocystis* or *Methylosinus* spp.

Screening for *pmoA*2 by Southern hybridization. In order to provide further evidence for the absence of *pmoA*2 in *M. trichosporium* strain OB3b and three type I MOB by a second, PCR-independent approach, a 26-bp oligonucleotide probe was formulated for the specific detection of *pmoA*2 (termed *pmoA*636b) (Table 2) by Southern hybridization. Based on the matching target region of *pmoA*2-like sequences initially available for the design of this probe, including environmental clone sequences directly retrieved from flooded rice microcosms, *Methylocystis* strain SC2 (one weak G:T mismatch) and *M. trichosporium* KS21 (Fig. 3), we concluded that *pmoA*636b should target a wide range of diverse *pmoA*2 sequence types (compare Fig. 1 and 3). The probe exhibited various mismatches to the corresponding target sites of *pmoA*1 of the MOB investigated by Southern hybridization. Southern blots of digested chromosomal DNA were hybridized, respectively, with the oligonucleotide probe *pmoA*636b and a *pmoA* gene probe (generated as a mixture of DIG-labeled *pmoA* amplicons of the different strains tested). All hybridization experiments were carried out with *Methylocystis* strain SC2 as a positive control. Under the hybridization and washing conditions used, probe *pmoA*636b did not detect *pmoA*1 for any of the MOB tested. As expected, no *pmoA*2 signal was observed for the type I MOB *M. capsulatus* strain Bath, *Methylocaldum* strain E10a, or *Methylomonas* strain D1a (Fig. 4). A *pmoA*2 signal was also not obtained for *M. trichosporium* strain OB3b, but the presence of *pmoA*2 was confirmed in another *M. trichosporium* strain (SM6) and in *Methylocystis* strain SC2. These results agreed well with those obtained by PCR. Full-length

gene probes of *pmoA* detected multiple *pmoA* copies in several of the strains. Three gene copies, including *pmoA*2, were detected in the genomic DNA of *M. trichosporium* strain SM6 and *Methylocystis* strain SC2, while two *pmoA* copies were observed for *M. trichosporium* strain OB3b and the three type I MOB (Fig. 4).

Expression of *pmoA*2 in *Methylocystis* strain SC2 (theoretical considerations). The paml model (35) was applied to a phylogenetic tree incorporating selected *pmoA*1 and *pmoA*2 gene sequences as shown in Fig. 5. The phylogeny was fit to a model incorporating eight separate ω values. All estimated ω values were <1 . The ω values within the *pmoA*1 and *pmoA*2 clusters were <0.1 , indicating that in recent evolutionary history, strong purifying selection has been acting on both genes. This is strong evidence that both genes are expressed and have important cellular functions. Interestingly, the branch connect-

Probe <i>pmoA</i> 636b	5' . . . CCGTGGGMCCTTATGTAGGCGTACTA . . . 3'
Target	
Clone M84-P3	3' . . . GGCACGCCCTGAATACATCCGCATGAT . . . 5'
Strain SC2, <i>pmoA</i> 2	. . . GGCACCCCGGAATACATCCGCATGGT . . .
Strain SC2, <i>pmoA</i> 1	. . . TCGAT GCCCGGAATATATCCGCATGAT . . .
Strain OB3b, <i>pmoA</i> 1	. . . TCGAT GCCCGGAATATATCCGCATGGT . . .
Strain SM6, <i>pmoA</i> 2	. . . GGCACGCCCGGAATACATCCGCATGGT . . .
Strain SM6, <i>pmoA</i> 1	. . . TCGAT GCCCGGAGTATATCCGCATGGT . . .
Strain D1a, <i>pmoA</i> 1	. . . GGTACACCTGAATACATCCGCATGGT . . .
Strain E10a, <i>pmoA</i> 1	. . . GGTACCCCGGAGTACATCCGCATGGT . . .
Strain Bath, <i>pmoA</i> 1	. . . GGTACGCCCTGAGTACATCCGCATGGT . . .

FIG. 3. Alignment showing the targeting region of probe *pmoA*636b among *pmoA*2 and corresponding *pmoA*1 sequences of those MOB investigated by Southern hybridization. The clone M84-P3 is shown as a representative of a set of diverse environmental *pmoA*2 sequences that perfectly match in the target region probe *pmoA*636b, including clones RsVc-03, Rr90a-43, RsVc-07, Rr90f-09, RsVc-05, and Rr90a-01 (Fig. 1). Letters shown in boldface indicate weak G:T mismatching nucleotide positions, while those shown in boldface and underlined indicate stronger mismatches (e.g., C:T, C:C, or G:G). Strains shown in the alignment are *Methylocystis* strain SC2, *M. trichosporium* strains OB3b and SM6, *Methylomonas* strain D1a, *Methylocaldum* strain E10a, and *M. capsulatus* strain Bath.

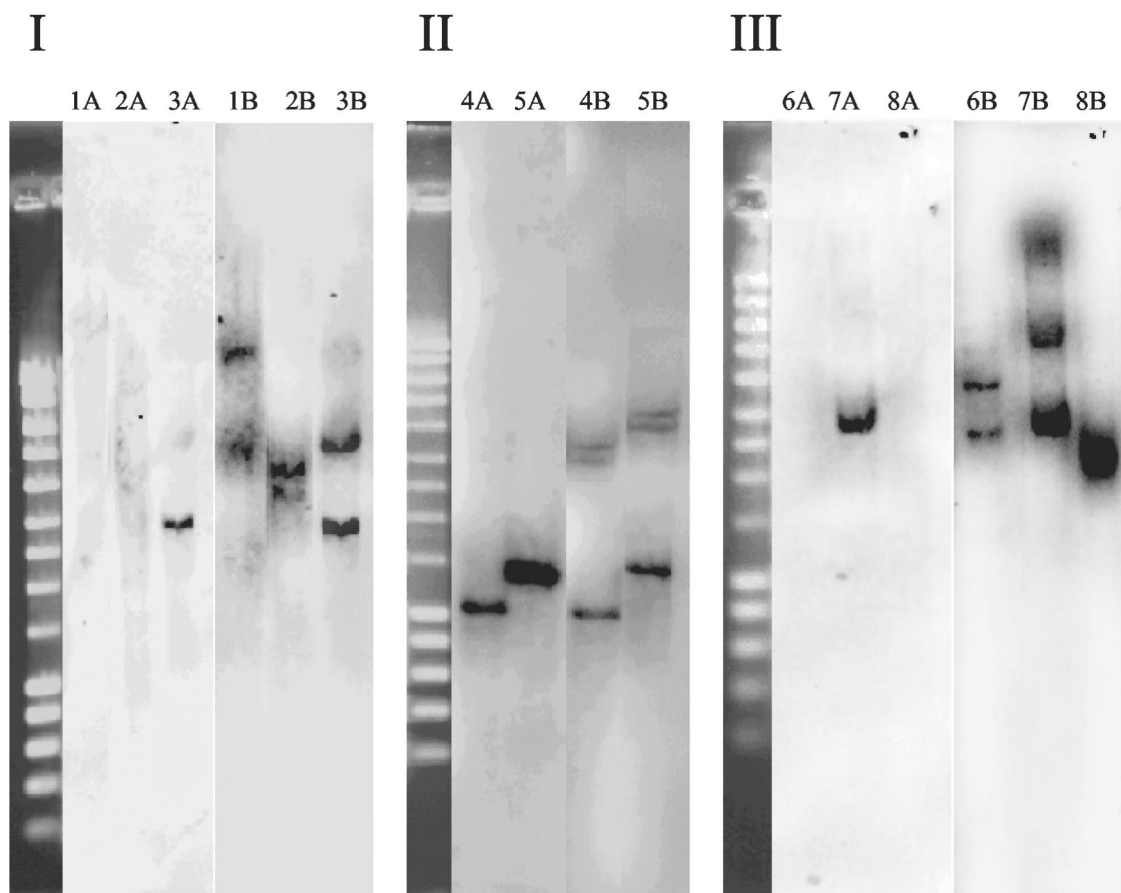


FIG. 4. Southern hybridization of genomic DNA to oligonucleotide probe pmoA636b (A lanes) or PCR-generated universal *pmoA* gene probe (B lanes). Under the hybridization conditions used in this study, probe pmoA636b is specific for *pmoA2* (Fig. 3), while the universal *pmoA* gene probe detects both *pmoA1* and *pmoA2*. (I) *Eco*RI digests. Lanes: 1, *M. trichosporium* strain OB3b; 2, *Methylocaldum* strain E10A; 3, *Methylocystis* strain SC2. (II) *Xho*I digests. Lanes: 4, *Methylosinus trichosporium* strain SM6; 5, *Methylocystis* strain SC2. (III) *Pst*I digests. Lanes: 6, *Methylococcus capsulatus* strain Bath; 7, *Methylocystis* strain SC2; 8, *Methylobomonas* strain D1a. *Methylocystis* strain SC2 was used as positive control (9). The *Pst*I digest of DNA of phage λ was used as a size marker.

ing the *pmoA2* cluster to the other branches of the phylogeny has an elevated ω value of 0.28. While still <1 , this increase may indicate a relaxation of purifying selection or even an increase in diversifying selection forces acting at particular times or on particular codons. Although purely speculative, a functional differentiation of the *pmoA1* and *pmoA2* genes may have occurred. Two different experimental strategies (i.e., Northern hybridization and RT-PCR) were applied to assess whether *pmoA2* was expressed when strain SC2 was grown under standard laboratory growth conditions (see Materials and Methods section "MOB strains").

Expression of *pmoA2* in *Methylocystis* strain SC2 (experimental assessment). Total RNA of *Methylocystis* strain SC2 was extracted and separated on a 1% formaldehyde agarose gel and blotted onto a nylon membrane for Northern hybridization. Probing was carried out with oligonucleotide A593b or pmoA636b specific for either *pmoA1* or *pmoA2*. In comparison to a strong signal observed for *pmoA1*, no conclusive signal was obtained for *pmoA2* (Fig. 6).

Consequently, we applied RT-PCR as a second, methodologically more sensitive, approach. Initially, we used the prim-

ers A593b and pmoA636b (Table 2) for RT. The primer combinations A189f-A593b and pmoA206f-pmoA636b were used to specifically amplify cDNA from either *pmoA1* or *pmoA2*. Sequencing of RT-PCR products confirmed strong expression of *pmoA1*. PCR of the cDNA generated with probe pmoA636b gave a product of nearly the size we expected. Sequencing of that product determined that it was not *pmoA*; it corresponded to a portion of 23S rRNA. We therefore assumed that non-specific cDNA was formed, perhaps because the primer may have bound to a nonspecific target at the lower temperature of RT. A similar problem occurred in our earlier attempts to perform RT-PCR (9) and may have contributed to our initial failure to detect *pmoA2* expression in strain SC2. In an attempt to overcome the problem, we used the universal *pmoA* primer A682b to generate the cDNA. Using the cDNA as template, we carried out specific PCR assays in which we applied either A593b or pmoA636b as the reverse primer. In both cases, products of the expected size were obtained (Fig. 7), and comparative sequence analysis of the RT-PCR product showed that sequences corresponded either to *pmoA1* or to *pmoA2*.

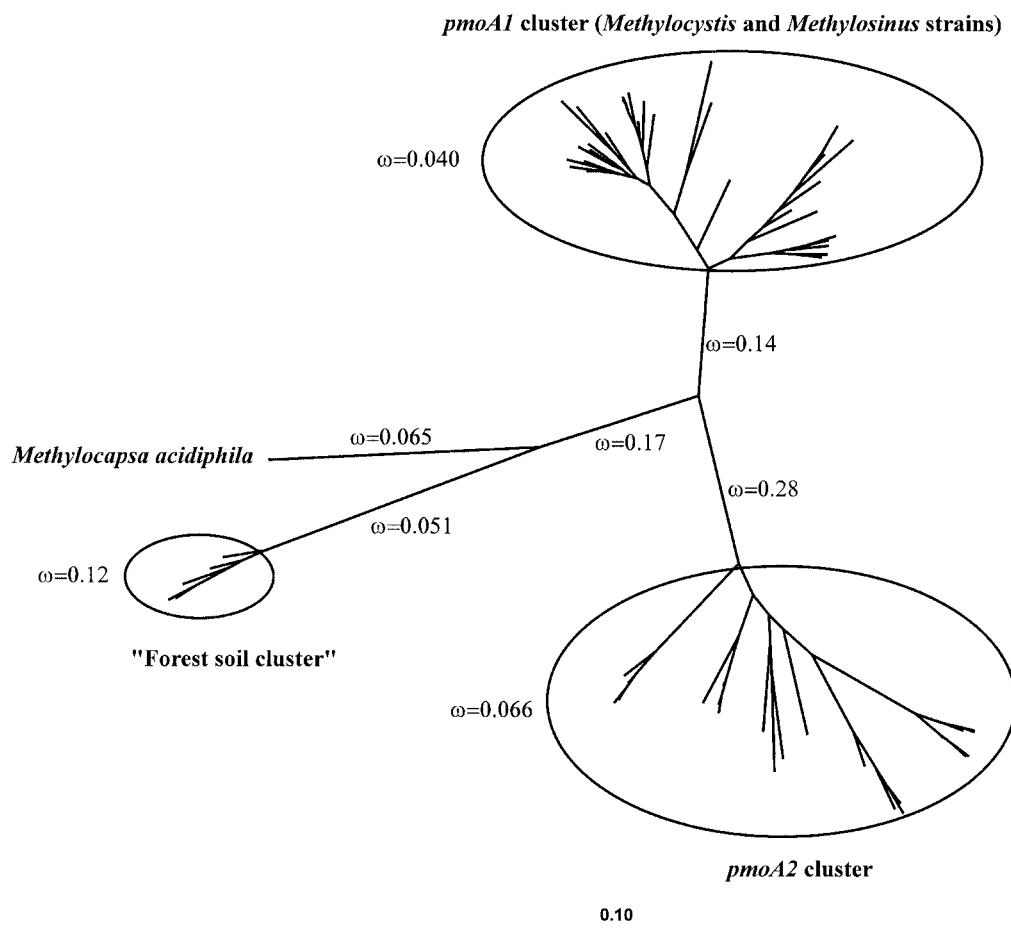


FIG. 5. Analysis of ω values within an unrooted phylogenetic tree incorporating 61 selected *pmoA* nucleotide sequences. The branching topology was calculated by using a neighbor-joining algorithm (29) with a Jukes-Cantor correction (20) (2,000 data resamplings), and branch lengths were adjusted by the paml program. A maximum-likelihood algorithm produced nearly the same tree topology. The distance bar indicates 0.1 nucleotide substitution per position. The paml model (model 2 in codeml) was fit with eight separate ω values as indicated: one for each of the five longest branches and one each for the *pmoA1* cluster, the *pmoA2* cluster, and the "forest soil cluster," respectively, which are enclosed by circles. The *pmoA1* cluster included sequences from various *Methylocystis* and *Methylosinus* strains (accession no. U31569, U31651, AJ431386, AJ431388, AJ458994, AJ458998, AJ459000-AJ459003, AJ459008, AJ459011, AJ459013, AJ459014, AJ459019, AJ459021-AJ459025, AJ459027-AJ459029, AJ459034, AJ459036, AJ459038, AJ459046-AJ459048, and AJ459052). The *pmoA2* cluster included sequences from pure cultures of *Methylocystis* and *Methylosinus* and sequences retrieved from soils by cultivation-independent methods (accession no. AJ431389, AJ299961, AY080950, and AJ459013 and sequences from the present study). The "forest soil cluster" contains sequences retrieved by Holmes et al. (18) using cultivation-independent methods (accession no. AF148525, AF148521, AF148527, AF148528, and AF148523).

DISCUSSION

The newly developed PCR assay enabled the differential detection of *pmoA1* and *pmoA2* in single MOB strains by T-RFLP analysis. The *pmoA*-based T-RFLP fingerprinting and comparative sequence analysis of 27 different type II MOB strains revealed that *pmoA2* is widely distributed among type II MOB, including strains of *Methylosinus sporium*, *M. trichosporium*, *Methylocystis echinoides*, and *M. parvus*, as well as a set of other *Methylocystis* strains (16). The MOB strains that tested positive for *pmoA2* can be considered to reflect the full phylogenetic diversity known within the *Methylosinus-Methylocystis* group. The treeing analysis clearly showed that *pmoA2* and deduced PmoA2 sequences form a coherent cluster clearly distinct from that of *pmoA1* and deduced PmoA1, respectively. It is noteworthy that PmoA1 sequences of the *Methylosinus-Methylocystis* group clustered in the maximum-likelihood tree

(Fig. 1) more closely to the PmoA sequence from *Methylocapsa acidiphila* than to their own corresponding PmoA2 sequences (Fig. 1).

The data obtained also suggested that *pmoA2* is not present in all type II MOB investigated. *M. trichosporium* strains OB3b and SM6 exhibit an overall 16S rRNA gene sequence identity of 98.2%. PCR-based screening led to the identification of *pmoA2* in strain SM6, but *pmoA2* could not be detected in strain OB3b. Further evidence that, in contrast to strain SM6, *M. trichosporium* strain OB3b does not harbor a *pmoA2* gene copy was provided by Southern hybridization using probe pmoA636b (Fig. 3 and 4). It should also be noted that the gene probe of PCR-amplified *pmoA* (containing both the conventional *pmoA1* and the novel *pmoA2*) led to the detection of two *pmoA* gene copies in *M. trichosporium* strain OB3b (Fig. 4). This is consistent with previous reports that strain OB3b har-

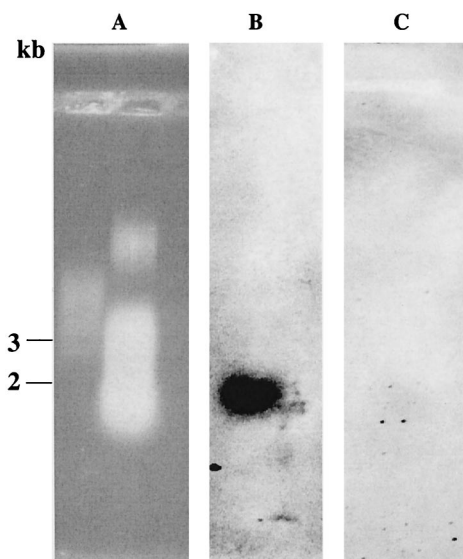


FIG. 6. Northern hybridization of total RNA from *Methylocystis* strain SC2 to *pmoA*-targeted oligonucleotide probes. (A) Agarose gel electrophoresis of total RNA in presence of an RNA marker (RNA ladder; New England Biolabs). (B) Hybridization with probe A593b to specifically detect mRNA expressed from *pmoA1*. (C) Hybridization with probe pmoA636b to specifically detect mRNA expressed from *pmoA2*. Lanes contain approximately 60 μ g of total RNA from strain SC2.

bors two nearly identical *pmoCAB* operons that correspond to *pmoA1* (11). Thus, the two different experimental approaches (PCR-based screening and Southern blot analysis) taken together provide strong evidence that closely related type II MOB genospecies may or may not harbor *pmoA2*. Interestingly, a similar observation has been made for the distribution of *mmoX* genes among type II MOB. Nearly identical species appear to differ in whether they possess or do not possess the *mmoX* gene (13, 16).

However, we cannot completely rule out the possibility

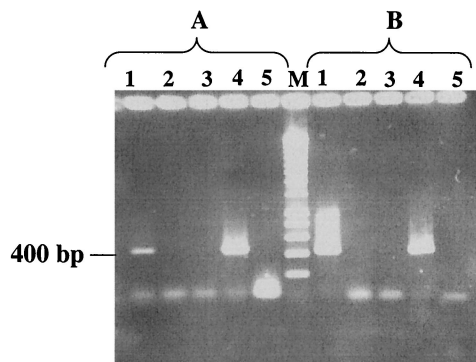


FIG. 7. Detection of *pmoA*-like mRNA in *Methylocystis* strain SC2 by RT-PCR. RT was carried out with primer A682b. (A) Use of the primer combination pmoA206f-pmoA636b for specific detection of functional activity of *pmoA2*. (B) Use of the primer combination A189f-A593b for specific detection of functional activity of *pmoA1*. Lanes: 1, total RNA plus RT-PCR (positive detection); 2, total RNA plus PCR without the RT step (negative control); 3, RNase-treated RNA sample plus PCR (negative control); 4, genomic DNA plus PCR (positive control); 5, no DNA plus PCR (negative control).

that variations in the target site of the primer combination pmoA206f-pmoA703b may be the reason why *pmoA2* was not detected in all type II MOB tested. Thus, the distribution of *pmoA2* among type II MOB might still have to be addressed in further studies. This may also be true for type I MOB, although in all five strains tested, the detection of a *pmoA2*-like gene copy by PCR (Table 1)—and in some cases by Southern hybridization (Fig. 4)—failed.

The second major aim of the present study was to determine whether *pmoA2* is expressed and encodes a functionally active protein or is instead an unexpressed pseudogene. Proteins with important basic cellular functions are often subject to strong purifying selection. Under such selection, a synonymous (non-amino acid changing) nucleotide mutation in the encoding gene has a much greater chance of becoming fixed than does a nonsynonymous (amino acid changing) mutation. Nonsynonymous/synonymous substitution rate ratios (ω) can typically be 0.1 or less for genes under strong purifying selection (36). Using the paml model (35), our phylogenetic analysis indicated that the *pmoA1* and *pmoA2* genes have both displayed similarly low ω values in recent evolutionary history. This is strong evidence that both genes are expressed and have important cellular functions. In contrast to our previous study (12), experimental evidence that *pmoA2* of *Methylocystis* strain SC2 is expressed could be obtained by RT-PCR (Fig. 7). However, as concluded from Northern hybridization experiments (Fig. 6), the level of *pmoA2* expression under the laboratory standard growth conditions applied was clearly lower than that of *pmoA1*.

Further studies will need to focus on the functional role of the enzyme expressed by the *pmo* operon that corresponds to *pmoA2*. Multiple pMMOs with different substrate specificities would provide MOB with a means of adapting to changing concentrations of methane and competitive substrates such as ammonia (13). There is some evidence for such a mechanism. Increasing Cu concentration has been shown to allow *M. capsulatus* to form a pMMO with a higher affinity for CH_4 (23). *Methylocystis* strain LR1, which possesses a *pmoA2*, displays a variable affinity for methane depending on the growth conditions (7), although this may be due to physiological conditions rather than to genetic changes (8). Ammonia does not appear to affect all soil MOB communities in the same way, suggesting that different MOB have different sensitivities to this cosubstrate (12). Thus, further investigation will require various cultivation conditions and the use of knockout mutations.

ACKNOWLEDGMENTS

This work was supported partly by grants from the Deutsche Forschungsgemeinschaft (SFB 395 to W. Liesack and grant DU377/1-1 to P. Dunfield) and Bundesministerium für Bildung und Forschung (bmb+f, GenoMik Network, Göttingen, Germany). M. Tchawa Yimga was supported by the German Academic Exchange Service (DAAD). J. Heyer was supported in part by the bmb+f BIOLOG project.

REFERENCES

- Adachi, J., and M. Hasegawa. 1996. Model of amino acid substitution in proteins encoded by mitochondrial DNA. *J. Mol. Evol.* **42**:459–468.
- Bowman, J. P., L. I. Sly, P. D. Nichols, and A. C. Hayward. 1993. Revised taxonomy of the methanotrophs: description of *Methylobacter* gen. nov., emendation of *Methylococcus*, validation of *Methylosinus* and *Methylocystis* species, and a proposal that the family *Methylococcaceae* includes only the group I methanotrophs. *Int. J. Syst. Bacteriol.* **43**:735–753.
- Brosius, J., M. L. Palmer, P. J. Kennedy, and H. R. Noller. 1978. Complete nucleotide sequence of a 16S ribosomal RNA gene from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **75**:4801–4805.

4. Conrad, R. 1996. Soil microorganisms as controllers of atmospheric trace gases (H_2 , CO, CH_4 , OCS, N_2O , and NO). *Microbiol. Rev.* **60**:609–640.
5. Dedysh, S. N., W. Liesack, V. N. Khmelenina, N. E. Suzina, Y. A. Trotsenko, J. D. Semrau, A. M. Bares, N. S. Panikov, and J. M. Tiedje. 2000. *Methylocella palustris* gen. nov., sp. nov., a new methane-oxidizing bacterium from peat bogs, representing a novel subtype of serine-pathway methanotrophs. *Int. J. Syst. Evol. Microbiol.* **50**:955–969.
6. Dedysh, S. N., V. N. Khmelenina, N. E. Suzina, Y. A. Trotsenko, J. D. Semrau, W. Liesack, and J. M. Tiedje. 2002. *Methylocapsa acidiphila* gen. nov., sp. nov., a novel methane-oxidizing and dinitrogen-fixing acidophilic bacterium from *Sphagnum* bog. *Int. J. Syst. Evol. Microbiol.* **52**:251–261.
7. Dunfield, P. F., W. Liesack, T. Henckel, R. Knowles, and R. Conrad. 1999. High-affinity methane oxidation by a soil enrichment culture containing a type II methanotroph. *Appl. Environ. Microbiol.* **65**:1009–1014.
8. Dunfield, P. F., and R. Conrad. 2000. Starvation alters the apparent half-saturation constant for methane in the type II methanotroph *Methylocystis* strain LR1. *Appl. Environ. Microbiol.* **66**:4136–4138.
9. Dunfield, P. F., M. Tchawa Yimga, S. N. Dedysh, U. Berger, W. Liesack, and J. Heyer. 2002. Isolation of a *Methylocystis* strain containing a novel *pmoA*-like gene. *FEMS Microbiol. Ecol.* **41**:17–26.
10. Felsenstein, J. 1993. PHYLIP (phylogeny inference package), version 3.5c. Department of Genetics, University of Washington, Seattle.
11. Gilbert, B., I. R. McDonald, R. Finch, G. P. Stafford, A. K. Nielsen, and J. C. Murrell. 2000. Molecular analysis of the *pmo* (particulate methane monooxygenase) operons from two type II methanotrophs. *Appl. Environ. Microbiol.* **66**:966–975.
12. Gullede, J., A. P. Doyle, and J. P. Schimel. 1997. Different NH_4^+ -inhibition patterns of soil CH_4 consumption: a result of distinct CH_4 -oxidizer populations across sites? *Soil Biol. Biochem.* **29**:13–21.
13. Hanson, R. S., and T. E. Hanson. 1996. Methanotrophic bacteria. *Microbiol. Rev.* **60**:439–471.
14. Henckel, T., M. Friedrich, and R. Conrad. 1999. Molecular analyses of the methane-oxidizing microbial community in rice field soil by targeting the genes of the 16S rRNA, particulate methane monooxygenase, and methanol dehydrogenase. *Appl. Environ. Microbiol.* **65**:1980–1990.
15. Heyer, J., Y. Malashenko, U. Berger, and E. Budkova. 1984. Verbreitung methanotropher Bakterien. *Z. Allg. Mikrobiol.* **24**:725–744.
16. Heyer, J., V. F. Galchenko, and P. F. Dunfield. 2002. Molecular phylogeny of type II methane-oxidizing bacteria isolated from various environments. *Microbiology* **148**:2831–2846.
17. Holmes, A. J., A. Costello, M. E. Lidstrom, and J. C. Murrell. 1995. Evidence that particulate methane monooxygenase and ammonia monooxygenase may be evolutionarily related. *FEMS Microbiol. Lett.* **132**:203–208.
18. Holmes, A. J., P. Roslev, I. R. McDonald, N. Iversen, K. Henriksen, and J. C. Murrell. 1999. Characterization of methanotrophic bacterial populations in soils showing atmospheric methane uptake. *Appl. Environ. Microbiol.* **65**:3312–3318.
19. Horz, H.-P., M. Tchawa Yimga, and W. Liesack. 2001. Detection of methanotroph diversity on roots of submerged rice plants by molecular retrieval of *pmoA*, *mmoX*, *mxsF*, and 16S rRNA and ribosomal DNA, including *pmoA*-based terminal restriction fragment length polymorphism profiling. *Appl. Environ. Microbiol.* **67**:4177–4185.
20. Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules, p. 21–132. In H. N. Munro (ed.), *Mammalian protein metabolism*. Academic Press, Inc., New York, N.Y.
21. Lane, D. J. 1991. 16S/23S rRNA sequencing, p. 115–175. In E. Stackebrandt and M. Goodfellow (ed.), *Nucleic acid techniques in bacterial systematics*. J. Wiley & Sons, Ltd., Chichester, United Kingdom.
22. Leadbetter, E. R., and J. W. Forster. 1958. Studies on some methane-utilising bacteria. *Arch. Microbiol.* **30**:91–118.
23. Lontoh, S., and J. D. Semrau. 1998. Methane and trichloroethylene degradation by *Methylosinus trichosporium* OB3b expressing particulate methane monooxygenase. *Appl. Environ. Microbiol.* **64**:1106–1114.
24. Lukow, T., P. F. Dunfield, and W. Liesack. 2000. Use of the T-RFLP technique to assess spatial and temporal changes in the bacterial community structure within an agricultural soil planted with transgenic and non-transgenic potato plants. *FEMS Microbiol. Ecol.* **32**:241–247.
25. Morris, S. A., S. Radajewski, T. W. Willison, and J. C. Murrell. 2002. Identification of the functionally active methanotroph population in a peat soil microcosm by stable-isotope probing. *Appl. Environ. Microbiol.* **68**:1446–1453.
26. Murrell, J. C., B. Gilbert, and I. R. McDonald. 2000. Molecular biology and regulation of methane monooxygenase. *Arch. Microbiol.* **173**:325–332.
27. Radajewski, S., G. Webster, D. S. Reay, S. A. Morris, P. Ineson, D. B. Nedwell, J. I. Prosser, and J. C. Murrell. 2002. Identification of active methylotroph populations in an acidic forest soil by stable-isotope probing. *Microbiology* **148**:2331–2342.
28. Reeburgh, W. S., S. C. Whalen, and M. J. Alperin. 1993. The role of methylotrophy in the global methane budget, p. 1–14. In J. C. Murrell and D. P. Kelly (ed.), *Microbial growth on C1 compounds*. Intercept Ltd., Andover, United Kingdom.
29. Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406–425.
30. Semrau, J. D., A. Chistoserdov, J. Lebron, A. Costello, J. Davagnino, E. Kenna, A. J. Holmes, R. Finch, J. C. Murrell, and M. E. Lidstrom. 1995. Particulate methane monooxygenase genes in methanotrophs. *J. Bacteriol.* **177**:3071–3079.
31. Stolyar, S., A. M. Costello, T. L. Peebles, and M. E. Lidstrom. 1999. Role of multiple gene copies in particulate methane monooxygenase activity in the methane-oxidizing bacterium *Methylococcus capsulatus* Bath. *Microbiology* **145**:1235–1244.
32. Strimmer, K., and A. von Haeseler. 1996. Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **13**:964–969.
33. Strimmer, K., N. Goldman, and A. von Haeseler. 1997. Bayesian probabilities and quartet puzzling. *Mol. Biol. Evol.* **14**:210–211.
34. Weisburg, W. G., S. M. Barns, D. A. Pelletier, and D. J. Lane. 1991. 16S ribosomal DNA amplification for phylogenetic study. *J. Bacteriol.* **173**:697–703.
35. Yang, Z., and J. P. Bielawski. 2000. Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.* **15**:496–503.
36. Yang, Z., R. Nielsen, N. Goldman, and A. M. K. Pedersen. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**:431–449.

3.1.2 Vergleichende Analyse der Operons von pMMO-1 und pMMO-2 von *Methylocystis* sp. Stamm SC2

Übersicht :

Die Erstellung einer genomischen BAC-Bibliothek des Typ II MB *Methylocystis* sp. Stamm SC2 ermöglichte die gezielte Sequenzierung der für die konventionelle bzw. die neuartige Kopie der pMMO kodierenden Gene. Erstmals konnte gezeigt werden, daß *pmoA2* von homologen Genen der in konventionellen pMMO-Operons vorhandenen Gene *pmoC* und *pmoB* flankiert wird. Die Gene *pmoC2*, *pmoA2* und *pmoB2* bilden ein Operon und werden gemeinsam transkribiert. Die vergleichende Analyse der erzielten DNA-Sequenzen von *pmoCAB1* und *pmoCAB2* resultierte in moderaten Identitätswerten. Die Primär- und insbesondere Sekundärstrukturen der abgeleiteten Polypeptide hingegen weisen deutlich höhere Ähnlichkeiten auf. Anhand der abgeleiteten Aminosäure-Sequenzen konnte nachgewiesen werden, daß es sich bei pMMO-2 um ein ebenfalls Membran-assoziiertes Enzym handelt, das viele der von pMMO-1 bekannten Eigenschaften aufweist. Obwohl pMMO-2 sich an 13 von 203 bei konventionellen pMMOs und AMOs hochkonservierten Positionen von deren Konsensussequenz unterscheidet, deuten die Untersuchungen auf eine Funktion der pMMO als Methan-Monooxygenase, möglicherweise mit einer abweichenden ökophysiologischen Rolle, hin.

Mittels der RT-PCR basierten 5'-RACE-Technik konnten die Transkriptions-Startpunkte beider Operons sowie entsprechende Promoter-Regionen festgestellt werden. Auch der Nachweis der Transkription von *pmoCAB1* und *pmoCAB2* war möglich.

Anmerkung :

Die "Shotgun"-basierte Sequenzierung des die *pmoCAB2* enthaltenden BAC-Klons wurde im Rahmen des Kompetenznetzwerks GenoMik am MPI für molekulare Genetik, Berlin von Dr. M. Kube durchgeführt

Comparative Analysis of the Conventional and Novel *pmo* (Particulate Methane Monooxygenase) Operons from *Methylocystis* Strain SC2

Peter Ricke,¹ Christoph Erkel,¹ Michael Kube,² Richard Reinhardt,² and Werner Liesack^{1*}

Max-Planck-Institut für terrestrische Mikrobiologie, D-35043 Marburg,¹ and Max-Planck-Institut für Molekulare Genetik, D-14195 Berlin-Dahlem,² Germany

Received 12 October 2003/Accepted 20 January 2004

In addition to the conventional *pmoA* gene (*pmoA1*) encoding the active site polypeptide of particulate methane monooxygenase, a novel *pmoA* gene copy (*pmoA2*) is widely distributed among type II methanotrophs (methane-oxidizing bacteria [MOB]) (M. Tchawa Yimga, P. F. Dunfield, P. Ricke, J. Heyer, and W. Liesack, Appl. Environ. Microbiol. 69:5593–5602, 2003). Here we report that the *pmoA1* and *pmoA2* gene copies in the type II MOB *Methylocystis* strain SC2 are each part of a complete *pmoCAB* gene cluster (*pmoCAB1*, *pmoCAB2*). A bacterial artificial chromosome (BAC) library of strain SC2 genomic DNA was constructed, and BAC clones carrying either *pmoCAB1* or *pmoCAB2* were identified. Comparative sequence analysis showed that these two gene clusters exhibit low levels of identity at both the DNA level (67.4 to 70.9%) and the derived protein level (59.3 to 65.6%). In contrast, the secondary structures predicted for PmoCAB1 and PmoCAB2, as well as the derived transmembrane-spanning regions, are nearly identical. This suggests that PmoCAB2 is, like PmoCAB1, a highly hydrophobic, membrane-associated protein. A total of 190 of the 203 amino acid residues representing a highly conserved consensus sequence of the currently known PmoCAB1 and AmoCAB sequence types could be identified in PmoCAB2. The *amoCAB* gene cluster encodes ammonia monooxygenase and is evolutionarily related to *pmoCAB*. Analysis of a set of amino acid residues that allowed differentiation between conventional *pmoA* and *AmoA* provided further support for the hypothesis that *pmoCAB2* encodes a functional equivalent of PmoCAB1. In experiments in which we used 5' rapid amplification of cDNA ends we identified transcriptional start sites 320 and 177 bp upstream of *pmoC1* and *pmoC2*, respectively. Immediately upstream of the transcriptional start sites of both *pmoCAB1* and *pmoCAB2*, sequence motifs similar to *Escherichia coli* σ^{70} promoters were identified.

Methane-oxidizing bacteria (MOB) (methanotrophs) are able to utilize methane (CH₄) as a sole source of carbon and energy for growth (15). These bacteria play an important role in the global methane cycle by oxidizing CH₄ released by methanogens in freshwater sediments and wetlands and thus mitigate the global warming effect of this greenhouse gas (6, 31). Phylogenies based on 16S rRNA genes show that MOB form distinct lineages in the gamma subclass of the class *Proteobacteria* (type I MOB) and the alpha subclass of the *Proteobacteria* (type II MOB) (3, 7, 8, 15, 18, 24). The two types of methanotrophs can be distinguished on the basis of biochemical and ultrastructural features (3, 15, 38).

The first step in CH₄ oxidation, the conversion of methane to methanol, is carried out by a methane monooxygenase (MMO). This enzyme exists in two forms, a particulate, membrane-associated form (pMMO) and a soluble form (sMMO). The two forms of the enzyme differ in structure, in kinetic properties, and in the range of substrates which are utilized (23). Only a restricted number of MOB species possess sMMO, while almost all MOB possess pMMO. The only MOB lacking pMMO are *Methylocella palustris* (8) and *Methylocella silvestris* (10). In MOB that harbor both forms of MMO, sMMO is synthesized under copper-deficient conditions, while in the presence of even a minuscule amount of available Cu(II)

(0.85 to 1.0 μ mol/g [dry weight] of cells) only pMMO is synthesized (15, 27).

The pMMO gene cluster consists of three consecutive open reading frames (*pmoC*, *pmoA*, and *pmoB*) in both type I MOB (32, 35) and type II MOB (14). The *pmo* genes from *Methylococcus capsulatus* Bath are transcribed into a single 3.3-kb polycistronic mRNA (27). PmoA is presumed to contain the active site because it has been shown to be specifically labeled by [¹⁴C]acetylene, a suicide substrate for MMO (30, 40).

The type I MOB *Methylococcus capsulatus* Bath and *Methylobacterium album* BG8 (32, 35), as well as the type II organisms *Methylosinus trichosporium* OB3b and *Methylocystis* sp. strain M (14), have been shown to contain duplicate copies of the *pmo* operon. The sequences of the duplicate *pmoCAB* gene clusters are nearly identical (e.g., there are 13 differences in 3,183 bp in *M. capsulatus* Bath).

However, the type II MOB *Methylocystis* strain SC2 has recently been shown to contain two very different *pmoA*-like genes (11). A 495-bp fragment of one gene (conventional *pmoA*, *pmoA1*) exhibited very high levels of sequence homology to *pmoA* genes of other type II MOB (encoding PmoA amino acid sequences identical to those of some other *Methylocystis* strains). The corresponding fragment of the second gene (novel *pmoA*, *pmoA2*) exhibited only 73% identity with *pmoA1* at the nucleotide level and 68.5% identity (83% similarity) at the deduced amino acid level. Genes closely related to *pmoA2* of strain SC2 are widely but not universally present in type II MOB (36). No *pmoA2*-like sequences were detected in five representative type I MOB tested. Comparative se-

* Corresponding author. Mailing address: Max-Planck-Institut für terrestrische Mikrobiologie, Karl-von-Frisch-Str., D-35043 Marburg, Germany. Phone: 49 (6421) 178 720. Fax: 49 (6421) 178 809. E-mail: liesack@staff.uni-marburg.de.

quence analysis showed that all *pmoA2*-like sequences formed a coherent cluster that is clearly distinct from *pmoA1* sequences of type I and type II MOB and from *amoA* sequences of the *Nitrosomonas-Nitrospira* group. Reverse transcription-PCR provided evidence that *pmoA2* was expressed in strain SC2 under standard laboratory growth conditions (36).

Here we show that both *pmoA1* and *pmoA2* are part of complete *pmoCAB* gene clusters in *Methylocystis* strain SC2. Although the deduced amino acid sequences of PmoCAB2 are very different from those of PmoCAB1, the putative secondary structure and regions of transmembrane-spanning helices seem to be highly conserved in the two PmoCAB variants. The biochemical equivalent of PmoCAB1 is the particulate methane hydroxylase (pMH), which is the main component of functionally active pMMO. The pMH complex consists of the following three subunits: α (45 kDa, PmoB1), β (27 kDa, PmoA1), and γ (23 kDa, PmoC1) (26, 40). The three polypeptides associate by noncovalent bonds and form a single complex with a stoichiometry of 1:1:1 ($\alpha\beta\gamma$) (23). Since our data were derived from an analysis of *pmo* genes, here we mainly refer to PmoCAB1 and PmoCAB2 rather than to pMH.

MATERIALS AND METHODS

Bacterial strains and growth conditions. The conditions used for growth of *Methylocystis* strain SC2 were adapted from the conditions described by Heyer et al. (18, 19). For extraction of high-molecular-weight (HMW) DNA, cells were grown in liquid cultures in medium 10 containing NaNO_3 instead of NH_4Cl as the nitrogen source. The cultures were grown for 3 to 5 days at 30°C under a headspace containing 20% (vol/vol) CH_4 , 5% (vol/vol) CO_2 , and 75% (vol/vol) air. Cells were pelleted by centrifugation at $6,000 \times g$ for 30 min at 4°C and washed once with washing buffer (WB) (20 mM Tris-HCl, 50 mM EDTA; pH 8.0). The cell biomass was immediately used for extraction of DNA. Cells of *Escherichia coli* strain DH10B were grown overnight in Luria-Bertani medium in a liquid culture at 37°C.

Preparation of HMW DNA. For extraction and further treatment of HMW DNA from strain SC2, whole cells were embedded in agarose at a final density of 3×10^9 cells ml^{-1} by using protocols described previously (12, 34). The embedded cells were treated with lysis buffer (100 mM Tris-HCl, 100 mM EDTA, 1.5 M NaCl, 1% [wt/vol] hexadecyltrimethylammonium bromide, 2% [wt/vol] sodium dodecyl sulfate; pH 8.0) (9) for 20 h at 37°C, rinsed with 25 ml of WB, and incubated in 5 ml of proteinase K reaction buffer (100 mM EDTA, 0.2% sodium deoxycholate, 1% sodium lauryl sarcosine, 1 mg of proteinase K per ml; pH 8.0) at 50°C for 14 h. To remove residual proteins, agar plugs were washed four times for 1 h in WB. The DNA-containing agarose plugs were stored in 0.5 M EDTA (pH 8.0) at 4°C until they were used.

Construction of BAC library. Cloning of HMW DNA into a bacterial artificial chromosome (BAC) vector (pIndigoBAC-5 Cloning-Ready; Epicentre, Madison, Wis.) required partial digestion with restriction endonuclease HindIII. To obtain the maximum percentage of genomic DNA fragments in the desired size range, a time series of partial restriction digestions (29) was performed for each extract of HMW DNA prior to final mass digestion. To enable subsequent agarose digestion, the center part of a 1% pulse-field agarose gel (Bio-Rad, Madison, Wis.) was replaced by 1% low-melting-point agarose (Bio-Rad). DNA size fractionation was performed in relation to a low-range PFG marker (New England Biolabs, Beverly, Mass.) with a contour-clamped homogeneous electric field apparatus (CHEF-DR III; Bio-Rad) by using three distinct separation steps, as described by Osogawa et al. (28).

Ethidium bromide-stained gel slices containing both electrophoresed HMW DNA and the PFG marker allowed us to cut slices from the unstained agarose gel containing genomic DNA that was approximately 50 to 150 kb long. HMW DNA was extracted immediately by agarose digestion with β -agarase I as described by the manufacturer (New England Biolabs). To prevent shearing of HMW DNA, subsequent steps were carried out without removing the digested agarose.

Eluted HMW DNA was ligated with 100 ng of pIndigoBAC-5. The molar ratio of insert to vector DNA was approximately 1:10. Ligation was performed in a 100- μl (total volume) mixture with 800 U of T4 ligase (New England Biolabs) for 24 h. The reaction was stopped by heat inactivation at 70°C for 30 min. The

ligation mixture was spotted onto the middle of a microdialysis filter (pore size, 0.025 μm ; Millipore, Bedford, Mass.) and was desalted by floating the filter on sterile 20% (vol/vol) polyethylene glycol for 3 h. Subsequently, *E. coli* cells were transformed with desalted DNA (ElectroMax DH10B competent cells; Invitrogen Life Technologies; Carlsbad, Calif.) at 2.5 kV, 25 μF , and 100 Ω in a precooled 1-mm electroporation cuvette. A total of 960 single clones were picked and transferred into 96-well microtiter plates.

PCR-based screening for BAC clones carrying *pmo* genes. For identification of clones carrying either *pmoA1* or *pmoA2*, 10- μl aliquots of BAC clone-positive *E. coli* cells, which were grown in 12 wells (one row) of a 96-well microtiter plate, were combined. An aliquot (20 μl) of each row pool was combined to generate plate pools. Pooled cells were lysed by boiling for 15 min, and the cell debris was subsequently pelleted by centrifugation with a microcentrifuge at $8,000 \times g$ for 5 min. Microtiter plates that contained at least one clone carrying target DNA were identified by PCR-based amplification of *pmoA* genes by using 1 μl of the supernatant of each plate pool as the template. Row pools of test-positive plates were screened in the same way. Finally, clones of single rows that tested positive were analyzed separately.

Primers A189f and 682b and an annealing temperature of 50°C were used to specifically detect *pmoA1*, while primers PmoA206f and PmoA703b and an annealing temperature of 66°C were used to screen for *pmoA2* (Table 1). Each reaction mixture contained 1 μl of template, 10 μl of $10\times$ reaction buffer, 1.5 mM Mg^{2+} , each deoxynucleoside triphosphate at a concentration of 200 μM , 2.5 U of *Taq* polymerase (Promega, Madison, Wis.), and each primer at a concentration of 0.25 μM . Amplifications were performed in 100- μl (total volume) mixtures in 0.2-ml reaction tubes by using a DNA thermal cycler (model 2400; PE Applied Biosystems, Weiterstadt, Germany). The thermal profile was as follows: initial denaturation for 3 min at 94°C, followed by 30 cycles consisting of denaturation at 94°C for 40 s, primer annealing for 45 s at the temperatures indicated above, and elongation at 72°C for 75 s. The final elongation step was extended to 7 min. Aliquots of the amplicons were checked by electrophoresis on 1% agarose gels. Amplicons of the expected size were purified and sequenced to verify identification of test-positive clones. To ensure that complete *pmo* operons were located on the selected clones, the 5' and 3' termini of the (cloned) inserts were sequenced by using vector-specific primers. The data obtained were compared with sequences of known *pmoB* and *pmoC* genes.

Two clones that carried either *pmoA1* (clone SC2-VII-C1) or *pmoA2* (clone SC2-IX-C7) were selected for further analysis.

DNA sequencing of BAC clones. The primary structures of the *pmoCAB1* (clone SC2-VII-C1) and *pmoCAB2* (clone SC2-IX-C7) gene clusters and their flanking regions were determined by direct sequence analysis of BAC DNA by using a primer walking approach and oligonucleotide primers designed for accurate sequencing. The known partial sequences of *pmoA1* and *pmoA2* (11) were used as the starting points to formulate novel oligonucleotide primers. BAC DNA was extracted with a QIAGEN Plasmid MAXI kit (Qiagen, Hilden, Germany) used according to the manufacturer's instructions. Each sequencing reaction mixture contained 13 μl (approximately 200 to 800 ng) of BAC DNA, 30 pmol of primer, and 6 μl of BigDye terminator mixture (PE Applied Biosystems). After initial denaturation at 94°C for 3 min, the following profile was repeated 99 times: rapid thermal ramp to 94°C and denaturation for 30 s, primer annealing at 50°C for 20 s, and elongation at 60°C for 4 min. After completion of the reaction, cycle sequencing products were purified with AutoSeq G-50 columns (Amersham Pharmacia Biotech Inc., Piscataway, N.J.). Sequences determined by primer walking were edited manually and assembled by using EditSeq and Seqman II (DNASTAR Inc., Madison, Wis.). In addition, the *pmoA2*-carrying clone SC2-IX-C7 was fully sequenced by a shotgun cloning approach with a minimal sequence quality of phred 30.

Isolation of total RNA. Cells of strain SC2 were grown in batch cultures to an optical density at 600 nm of 0.5 to 0.7 (mid-exponential growth phase). Expression of *pmo* genes was promoted by ensuring that the gas phase contained 10% CH_4 approximately 20 h before RNA extraction. Total RNA was stabilized before cell lysis by application of the RNeasy Protect bacterial reagent (Qiagen). Aliquots containing 1×10^9 cells were lysed by mechanical disruption (bead beating at 2,500 rpm for 90 s), and RNA was extracted with an RNeasy mini kit (Qiagen). The manufacturer's protocol was modified slightly. Instead of the recommended on-column DNase digestion, which failed to remove DNA completely, the RNA preparations were treated with RNase-free DNase (RQ1; Promega) for 40 min at 37°C. The concentration of nucleic acids was determined by photometric measurement at 260 nm.

Identification of transcriptional start sites of *pmoCAB* gene clusters. cDNA was produced by using the ThermoScript reverse transcription-PCR system (Invitrogen) and a primer (Table 1) with target specificity for either *pmoC1* (ConvC100r) or *pmoC2* (NovC313r). ConvC100r and NovC313r were selected

TABLE 1. Oligonucleotide primers used in this study

Primer	Position	Sequence (5'→3') ^a	Direction
Primers for screening			
A189f ^b	1486–1504 ^c	GGN GAC TGG GAC TTC TGG	Forward
PmoA206f ^d	1486–1521 ^c	GGNGACTGGGACTTCTGGATCGACTTCAAGGATCG	Forward
682b ^d	2026–2009 ^c	GAA SGC NGA GAA GAA SGC	Reverse
PmoA703b ^b	2026–1992 ^c	GAASGCNGAGAAGAASGCGGCGACCGGAACGACGT	Reverse
RACE primers			
NovC40r	236–216 ^c	GAT CGT ATC GGT CTC GCC GGC	Reverse
NovC100r	293–273 ^c	GCC GAC GTA GAA GAC CAG CAT	Reverse
NovC313r	511–492 ^c	CAG CCG ATG AGC GTC AGG AT	Reverse
ConvC-60r	269–289 ^c	TGC GCG TGA CCG GTA TCG CTC	Reverse
ConvC40r	372–352 ^c	CAG CTT CCG TGC CAG CAG CAG	Reverse
ConvC100r	445–425 ^c	TCG TAA ATG CGA ACG ACC AGA	Reverse
RACE kit primers^f			
Oligod(T)Anchored	–38–1 ^{c,e}	GAC CAC GCG TAT CGA TGT CGA C(T) ₁₆ V	Forward
Anchor primer	–38–17 ^{c,e}	GAC CAC GCG TAT CGA TGT CGA C	Forward

^a N = A, C, G, or T; S = C or G; V = A, C, or G.

^b Data from reference 20.

^c Primer position in relation to the transcriptional start site of *pmoCAB1*.

^d Data from reference 36.

^e Primer position in relation to the transcriptional start site of *pmoCAB2*.

^f Obtained from Roche Diagnostics (Mannheim, Germany).

for analysis from a set of newly designed primers because they exhibited the best performance in cDNA synthesis. Aliquots (0.5 to 1 µg) of DNA-free RNA were used as templates and were transcribed at 56°C for 1 h. The cDNA was purified with a Qiaquick PCR purification kit (Qiagen).

The transcriptional start sites were determined by rapid amplification of cDNA ends (RACE) by using a 5'/3'RACE kit (Roche, Mannheim, Germany) according to the manufacturer's protocol. The oligonucleotide primers used are shown in Table 1. Briefly, a poly(A) tail was synthesized on the 3' end of the first-strand cDNA by using terminal transferase. The use of an anchored oligo(dT) primer and a primer targeting a region of *pmoC* (for *pmoC1*, primer ConvC40r; for *pmoC2*, primer NovC100r) allowed us to amplify tailed cDNA.

A second PCR was performed with a nested *pmoC*-targeting primer (for *pmoC1*, primer ConvC-60r; for *pmoC2*, primer NovC40r) and an anchored forward primer; a 1:50 dilution of the first PCR product was used as the template. First- and second-round PCR were carried out as described above with an annealing temperature of 61°C. PCR products were purified with a QIAquick PCR purification kit and were sequenced by using the BigDye terminator mixture (PE Applied Biosystems) as described by the manufacturer. The starting point of the poly(A) sequence indicated the 3' end of the cDNA, which corresponded to the putative transcriptional start site.

Computational analysis. Homology searches were carried out with standard software tools and the BLAST server of the National Center for Biotechnology Information by using the BLAST algorithm (1) with a nonredundant DNA database, as well as with a nonredundant protein database compiled for this study from SWISSPROT, TrEMBL, and PIR.

Putative factor-independent terminators were calculated by using the program TERMINATOR of the Genetics Computer Group software package (4). Similarity and identity values for deduced amino acid sequences were calculated by using the stretcher program (<http://bioweb.pasteur.fr/seqanal/interfaces/stretcher.html>) based on global pairwise alignments (25) computed with a BLOSUM65 (17) substitution matrix. Alternatively, similarity and identity values were calculated with the ARB program package (developed by O. Strunk and W. Ludwig of the Technical University of Munich [<http://www.arb-home.de>]), based on manually edited multiple alignments.

Hydrophobicity plots of PmoCAB1 and PmoCAB2 were calculated by using the comparison tool of the Protein Hydrophilicity/Hydrophobicity Search and Comparison Server (<http://bioinformatics.weizmann.ac.il/hydroph/>) of the Weizmann Institute of Science, Rehovot, Israel. The locations of putative transmembrane-spanning regions were computed by using the toppred transmembrane topology prediction program (<http://bioweb.pasteur.fr/seqanal/interfaces/toppred.html>).

Nucleotide sequence accession numbers. The sequences of the *pmoCAB* gene clusters plus flanking regions from *Methylocystis* strain SC2 have been deposited in the EMBL, DDBJ, and GenBank databases under accession numbers AJ584611 and BX649604 for *pmoCAB1* and *pmoCAB2*, respectively. The information on the flanking regions of *pmoCAB2* includes the annotation results for the fully sequenced 18.3-kb insert of clone SC2-IIX-C7.

RESULTS AND DISCUSSION

Methylocystis strain SC2 BAC clones carrying *pmo* genes.

Gilbert et al. (14) reported that in accordance with previous studies (32), cloning of *pmo* operons in multicopy vectors, such as pUC plasmids, is possible only in overlapping fragments because parts of the genes might be toxic to *E. coli*. Analysis of a *pmo* operon sequence by cloning of overlapping fragments involves the risk that sequence data obtained originate from different, multiple *pmo* operons. This problem is negligible if the sequences of multiple *pmo* operons are virtually identical, as is the case for the duplicate *pmo* operons of *M. trichosporium* OB3b (14). However, reconstruction of the primary operon structure from a genome containing multiple *pmo* operons with divergent sequences could be much more problematic. Thus, in order to avoid cloning of overlapping *pmo* fragments and to obtain full-length *pmo* operons on the cloned fragments, we used a single-copy BAC vector to construct a genomic library of strain SC2.

Using a coordinated PCR-based screening approach, we identified six and seven clones carrying *pmoA1* and *pmoA2*, respectively, in a set containing 960 BAC clones. To ensure that complete *pmo* operons were located on the BAC clones, the sequences of the terminal regions of the inserts were compared to known *pmoC* and *pmoB* sequences. One clone that carried *pmoCAB1* (clone SC2-VII-C1) and another clone that carried *pmoCAB2* (clone SC2-IIX-C7) were selected for further analysis. The complete sequences of the two *pmo* gene clusters plus their flanking regions were determined by a primer walking strategy. In addition, the complete sequence of the *pmoCAB2* cluster was confirmed by applying shotgun sequencing to clone SC2-IIX-C7. Comparison of the sequences determined for the *pmoCAB2* operon and flanking regions either by primer walking or by shotgun cloning did not reveal any ambiguous nucleotide positions.

Structural organization of *pmoCAB1* and *pmoCAB2*. Comparative sequence analysis revealed that *pmoA1* was present in a defined *pmoCAB* gene cluster. This gene arrangement agrees

TABLE 2. Comparison of *pmo* genes at the DNA and amino acids levels

Organisms	<i>pmoC</i>		<i>pmoA</i>		<i>pmoB</i>	
	% DNA identity	% Amino acid similarity (identity)	% DNA identity	% Amino acid similarity (identity)	% DNA identity	% Amino acid similarity (identity)
Strain SC2 <i>pmo2</i> -strain SC2 <i>pmo1</i>	69.0	74.8 (59.3)	70.9	78.9 (65.6)	67.4	75.4 (63.5)
Strain SC2 <i>pmo2</i> - <i>Methylocystis</i> sp. strain M	71.7	77.1 (64.0)	71.1	79.7 (65.2)	67.0	74.7 (61.4)
Strain SC2 <i>pmo2</i> - <i>M. trichosporium</i> OB3b	71.6	77.1 (63.6)	70.7	80.1 (64.5)	69.2	75.8 (64.1)
Strain SC2 <i>pmo2</i> - <i>M. capsulatus</i> Bath	67.5	71.6 (57.6)	65.5	74.6 (57.0)	60.5	67.1 (46.9)
Strain SC2 <i>pmo1</i> - <i>Methylocystis</i> sp. strain M	83.9	91.8 (85.9)	91.7	97.6 (95.2)	86.1	94.5 (89.5)
Strain SC2 <i>pmo1</i> - <i>M. trichosporium</i> OB3b	82.8	91.0 (80.5)	85.2	94.0 (87.3)	78.8	88.9 (81.2)
Strain SC2 <i>pmo1</i> - <i>M. capsulatus</i> Bath	68.8	72.9 (59.2)	66.6	76.2 (58.3)	59.2	61.2 (45.6)
<i>Methylocystis</i> sp. strain M- <i>M. trichosporium</i> OB3b	85.5	93.0 (85.9)	84.2	94.4 (86.5)	80.5	86.8 (78.7)
<i>Methylocystis</i> sp. strain M- <i>M. capsulatus</i> Bath	78.7	74.4 (60.3)	66.4	75.4 (57.5)	61.2	60.9 (45.2)
<i>M. trichosporium</i> OB3b- <i>M. capsulatus</i> Bath	69.2	73.3 (69.2)	67.4	76.2 (57.9)	60.6	61.8 (44.0)

amino acids were included in a further comparative analysis. The *pmo* operons determined for *M. trichosporium* OB3b and *Methylocystis* sp. strain M (14), as well as one of the two nearly identical *pmo* operons of the type I MOB *M. capsulatus* Bath (32, 35), were used for comparison (Table 2). Comparative analysis of the amino acid sequences deduced from *pmoC* and *pmoB* revealed that PmoC1 and PmoC2, as well as PmoB1 and PmoB2, have identity and similarity values which are in the same range as the values for partial PmoA1 and PmoA2 sequences (11). The levels of identity between polypeptides predicted for PmoCAB2 of strain SC2 and polypeptides predicted for the conventional PmoCAB proteins of type II MOB ranged from 59.3 to 65.6%, while the levels of similarity were in the range from 74.7 to 80.1%. Comparison of PmoCAB2 with PmoCAB of the type I MOB *M. capsulatus* Bath resulted in moderately lower levels of identity and similarity for PmoC and PmoA and clearly lower values for PmoB.

(ii) **Putative secondary structures.** The polypeptides predicted for both PmoCAB1 and PmoCAB2 are highly hydrophobic and contain several transmembrane-spanning regions (Fig. 2 and 3). All N-terminal sequences were predicted to be located in the cytosol. Comparison with putative transmembrane helices of PmoCAB subunits from *M. trichosporium* OB3b, *Methylocystis* sp. strain M, and *M. capsulatus* Bath revealed very high levels of structural similarity. The high degree of structural conservation of PmoCAB2 suggests that it clearly resembles PmoCAB1 in terms of its physicochemical properties. This allowed us to conclude that, like *pmoCAB1*, *pmoCAB2* encodes a membrane-associated protein (Fig. 3). As described by Gilbert and coworkers (14), the first PmoB transmembrane-spanning region (PmoB [helix 1]) (Fig. 3) has to be regarded hypothetically only, because the N-terminal region of PmoB is considered to constitute a leader sequence (26, 37, 40).

(iii) **Highly conserved amino acid residues.** Recently, Tukhvatullin et al. (37) described a structural and functional model for pMH from *M. capsulatus* Bath. This model was predicted based on amino acid residues found to be almost completely conserved in a data set consisting of 3 sequence types for PmoB1 and 12 sequence types for AmoB, 112 sequence types for PmoA1 and 349 sequence types for AmoA, and 5 sequence types for PmoC1 and 5 sequence types for AmoC. Because of the high degree of conservation, these amino acid positions are

believed to include various functionally important residues. Of the 85 residues determined by Tukhvatullin et al. (37) to be highly conserved in PmoB1 and AmoB (α -peptide), 79 were also identified in PmoB2. Likewise, PmoA2 (β -peptide) and PmoC2 (γ -peptide) possessed 34 of 36 and 77 of 82 highly conserved residues, respectively.

In total, the PmoCAB2 sequence had 13 deviations from the consensus sequence. PmoCAB1 of strain SC2 differed at only two positions (P86→A97 and I249→L254 of PmoC1; the numbering of the consensus sequence is that of Tukhvatullin et al. [37], who did not consider putative leader sequences). However, it should be noted that of the 13 residues at which PmoCAB2 differed from the PmoCAB1-AmoCAB consensus sequence, only a single residue is thought to have a different chemical property than the corresponding residue of the consensus sequence (R180 of PmoC2, compared with the consensus residue G167 in PmoC1 of *M. capsulatus* Bath). Each of the other 12 residues belonged to the same amino acid similarity group (as defined empirically by the BLOSUM65 matrix [17]) as its counterpart in the consensus sequence. Thus, the similarity value for the consensus sequence of PmoCAB1 and AmoCAB and the corresponding positions in PmoCAB2 (>99%) provides support for the assumption that *pmoCAB2* encodes a pMMO-like (or AMO-like) enzyme.

While the structure of pMMO and its content of metal centers is still under discussion (2, 5, 22), it is widely accepted that pMMO requires copper as a cofactor for both functional activity and structural integrity (2, 5). Tukhvatullin et al. (37) analyzed the set of highly conserved amino acid residues described above to examine their possible roles in the formation of complexes with metal ions. This analysis took into account the physicochemical characteristics of particular amino acid residues, as well as experimental data for pMMO and pMH obtained by electron spin resonance, electron nuclear double resonance, Mössbauer, and atomic absorption spectroscopy. The analysis aimed to elucidate elements of pMMO-AMO structure and to reveal amino acid residues which are involved in the formation of the active sites. Thirty-nine residues which have the potential to act as ligands were identified. Thirty-eight of these residues could also be identified in PmoCAB2. The only exception, Y26 (PmoA1 and AmoA), is replaced in PmoA2 by F (PmoA2 in Fig. 2 at position F44). Y26 is thought to be involved in formation of the catalytic center only in a

10	20	30	40	50	60
MQQTTEKAV <u>S</u>	AAQAGETDTI	V <u>N</u> YKPAYITM	GVMLVFYVGI	RVYEQYFGWK	AGLDSFAPEF
70	80	90	100	110	120
QTYWMNLMWT	ELPLEFVAF <u>C</u>	GIGGYLW <u>K</u> TR	DRNIDAVAPR	EEMRRILTLI	CWLAVYAFSV
130	140	150	160	170	180
YWGAS <u>Y</u> FT <u>E</u> Q	DGTWHQTVIR	DTDFTPSHIL	EFYLSYPIYI	ICGWGAFMYA	HTRIPQFA <u>K</u> R
190	200	210	220	230	240
ISLAFLMFFA	GPFMIFPN <u>I</u> G	LNEWGHTFWF	MEELFTAPLH	WGFVFFGWFA	LAVFGVARQV
250	258				
LDRVIELSKE	YEKDAL <u>A</u> L				
PmoC					
10	20	30	40	50	60
MSASIEGTSP	TGKTWKSKEE	FLGCVILT <u>D</u> W	ILLVILFAVL	LGS <u>F</u> HIHYML	LAGDWD <u>F</u> WID
70	80	90	100	110	120
FKDRRMWPTV	TPIVAMCFAA	AAQSFFWTRF	RLPIGATTVV	LALLIGEWI <u>N</u>	RYDNF <u>W</u> GWTF
130	140	150	160	170	180
FPINLVFP <u>S</u> A	LIPMGFWLDI	VLMISGSWL <u>V</u>	TALLGGLGWG	LLFY <u>P</u> INWPV	LAQYHQAAEI
190	200	210	220	230	240
DGVLLTLADL	IGFN <u>Y</u> VRTGT	PEYIRMVERG	TLRTFGKDVV	PVAAFFSGFI	SMLVYFLW <u>N</u>
250	256				
MGKWFSTTKY	LEIEEV				
PmoA					
10	20	30	40	50	60
MKSLNLHSM <u>A</u>	KCATARVVRL	WILGLAVAGA	LTTLPAATPA	AAHGERSQQA	FLRMRTLN <u>W</u> Y
70	80	90	100	110	120
DVKWSKTEL <u>N</u>	VNDEMELTGK	VHVFSGW <u>P</u> QA	VARPGESFL <u>N</u>	VGEPGPVLIR	KSAFVGEVPV
130	140	150	160	170	180
PRT <u>F</u> SMDVGN	DY <u>E</u> YKIVLKA	RRQGRYH <u>V</u> HV	QINVKDGGPI	VGPGQWITIK	GDMKDFTN <u>P</u> V
190	200	210	220	230	240
TLLEGSTIDL	ETYGITWTYI	<u>Y</u> HFFWMAAAA	AWILYWFTMK	GIIVRYWQVK	AGKGRELIGQ
250	260	270	280	290	300
DEKRIGAVTL	AAVLLAVLVF	YASANREFPR	TLPMQAGLLT	GISPIDEPRS	TVAVQPKGGS
310	320	330	340	350	360
YTVPGRELRI	RAKVTTNNGKE	PIRLGEFTSA	GLRFLNPVDF	TTRPDFDYL	LADRGLSVSD
370	380	390	400	410	420
PNPIAPGETR	DLEVIVQDAR	<u>F</u> DIERLSDLA	YDTSQFGGL	LFFFSPSGER	QRAEIGGPVI
429					
PKFQAGATL					
PmoB					

FIG. 2. Primary structures of derived PmoCAB2 expressed in one-letter code. Predicted transmembrane helices are indicated by shading. Residues that are highly conserved in PmoCAB1 and AmoCAB (consensus sequence) are indicated by boldface type (37). Amino acids that are located at conserved positions but differ from the amino acids in the consensus sequence determined by Tukhvatullin et al. (37) are underlined. The N-terminal helix of PmoB, as predicted by toppred and shown in Fig. 3, is not indicated because the residues are thought to constitute a leader sequence (26, 37).

subset of the structural-functional models predicted. However, in the majority of suggested models, Y26 is replaced as a ligand by other residues. Thus, the replacement of tyrosine by phenylalanine does not necessarily have an impact on the functional properties of PmoA2.

Among the PmoCAB1 and AmoCAB sequences, the largest data set is available for PmoA1 and AmoA, which include the active sites of the enzymes. To define the putative cellular function of PmoCAB2 more precisely, we identified in 315 PmoA1 and 919 AmoA sequences a set of 18 amino acid positions that can be used to differentiate between conventional pMMO and AMO (Table 3). Analysis of these signature positions in 40 PmoA2 sequences clearly showed that PmoA2 agrees with PmoA1 at all but one position.

Taken together, the analysis of a set of amino acid residues that are highly conserved among all pMMO and AMO and of a set that can be used to differentiate between the two enzymes

clearly suggests that *pmoCAB2* encodes a functional equivalent to conventional pMMO. Calculation of the ratio of nonsynonymous nucleotide substitutions to synonymous nucleotide substitutions in a phylogeny (36, 39) suggested that in recent evolutionary history *pmoA2* was subject to strong purifying selection and therefore has an important cellular function. However, the analyses also suggested that *pmoA2* may have been subject to diversifying selection forces acting at particular times or on particular codons. The elevated number of deviations in the PmoCAB2 sequence from the consensus sequence of PmoCAB1 and AmoCAB supports the hypothesis that there have been diversifying selection forces and may suggest that the novel pMMO-like enzyme has a different functional role in strain SC2 than the conventional pMMO has. To obtain further evidence for this hypothesis, we are currently assessing in cultivation studies the response of the level of *pmoCAB2* expression to various environmental conditions (e.g., different concentrations of methane and competitive

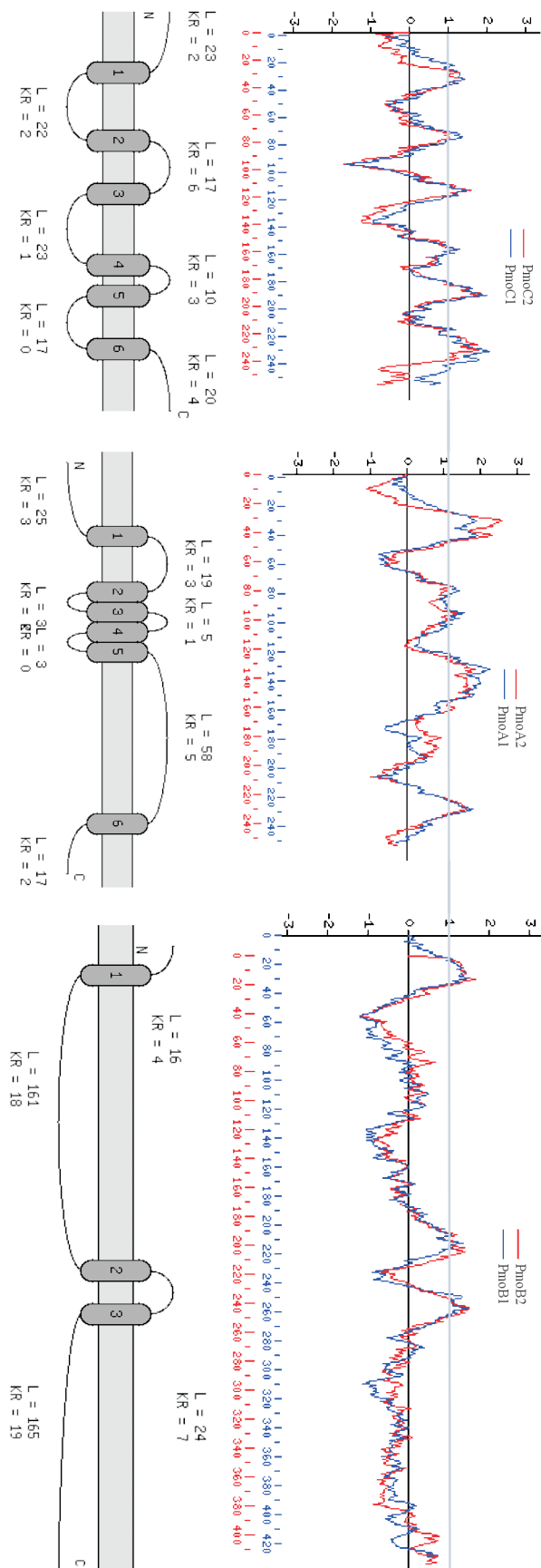


FIG. 3. Predicted topologies of derived PmoCAB1 and PmoCAB2 from *Methylocystis* strain SC2. (Top) Hydrophobicity plot of PmoCAB1 versus PmoCAB2. The y axis indicates the relative hydrophobicity value at a given position in PmoCAB. (Bottom) Schematic presentation of predicted transmembrane helices of PmoCAB2. N-terminal sequences are predicted to be located in the cytosol. Exact locations of transmembrane-spanning regions are shown in Fig. 2.

TABLE 3. Amino acid residues enabling differentiation of AmoA, PmoA1, and PmoA2

Position ^a	Residue in:		
	AmoA ^b	PmoA1	PmoA2
142	M ^c	I	L
145	T	S	S
147	n (d) ^d	S (A)	S (A)
157 ^e	(a)	(i)	(I)
171	F (F)	I (I)	I (i)
172	G	a	A
179	V (I)	e (d)	E (D)
185	L	m	I
190	(f)	(i)	(I)
204	v	i	I
206	l	m	M
207	I	v	v (I)
209	(d)	(h)	(H)
211	S	t	T
217	G	k	K
218	H (H)	d (d)	D (D)
219	T (A)	v (i)	V (I)
221	V (I)	p (a)	P (A)

^a The positions are the positions in PmoA2, as shown in Fig. 2.

^b Identification of signature residues indicative of either AmoA or PmoA was based on 919 AmoA sequences, 315 PmoA1 sequences, and 40 PmoA2 sequences.

^c Amino acid residues are indicated by the one-letter codes. Uppercase letters indicate residues which are conserved in at least 95% of the reference data set, while lowercase letters indicate residues which are conserved in at least 80% of the sequences compared.

^d The letters in parentheses represent amino acid similarity groups, as follows: A, PAGST; D, QNEDBZ; H, HKR; I, LIVM; and F, FYW.

^e A highly conserved amino acid residue could not be identified for positions 157, 190, and 209. However, the majority of residues identified for each individual data set at these positions belonged to the same amino acid similarity group.

substrates) and trying to generate mutants defective in either *pmoCAB1* or *pmoCAB2*. The *pmoCAB2* operon is expressed under standard laboratory growth conditions (36; this study), but its level of expression is clearly lower than that of *pmoCAB1* (36). This observation may be a further indication that the novel enzyme has a different functional role than the conventional pMMO has.

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (grant SFB 395 to W. Liesack) and by the Competence Network Göttingen project Genome Research on Bacteria (GenoMik) financed by the German Federal Ministry of Education and Research (BMBF).

REFERENCES

- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410.
- Basu, P., B. Katterle, K. K. Andersson, and H. Dalton. 2003. The membrane-associated form of methane mono-oxygenase from *Methylococcus capsulatus* (Bath) is a copper/iron protein. *Biochem. J.* **369**:417–427.
- Bowman, J. P., L. Jimenez, I. Rosario, T. C. Hazen, and G. S. Saylor. 1993. Characterization of the methanotrophic bacterial community present in a trichloroethylene-contaminated subsurface groundwater site. *Appl. Environ. Microbiol.* **59**:2380–2387.
- Brendel, V., and E. N. Trifonov. 1984. A computer algorithm for testing potential prokaryotic terminators. *Nucleic Acids Res.* **12**:4411–4427.
- Choi, D. W., R. C. Kunz, E. S. Boyd, J. D. Semrau, W. E. Antholine, J. I. Han, J. A. Zahn, J. M. Boyd, A. M. de la Mora, and A. A. DiSpirito. 2003. The membrane-associated methane monooxygenase (pMMO) and pMMO-NADH:quinone oxidoreductase complex from *Methylococcus capsulatus* Bath. *J. Bacteriol.* **185**:5755–5764.
- Conrad, R. 1996. Soil microorganisms as controllers of atmospheric trace gases (H₂, CO, CH₄, OCS, N₂O, and NO). *Microbiol. Rev.* **60**:609–640.
- Dedysh, S. N., V. N. Khmelenina, N. E. Suzina, Y. A. Trotsenko, J. D. Semrau, W. Liesack, and J. M. Tiedje. 2002. *Methylocapsa acidiphila* gen. nov., sp. nov., a novel methane-oxidizing and dinitrogen-fixing acidophilic bacterium from *Sphagnum* bogs. *Int. J. Syst. Evol. Microbiol.* **52**:251–261.
- Dedysh, S. N., W. Liesack, V. N. Khmelenina, N. E. Suzina, Y. A. Trotsenko, J. D. Semrau, A. M. Bares, N. S. Panikov, and J. M. Tiedje. 2000. *Methylocella palustris* gen. nov., sp. nov., a new methane-oxidizing acidophilic bacterium from peat bogs, representing a novel subtype of serine-pathway methanotrophs. *Int. J. Syst. Evol. Microbiol.* **50**:955–969.
- Dedysh, S. N., N. S. Panikov, W. Liesack, R. Großkopf, J. Zhou, and J. M. Tiedje. 1998. Isolation of acidophilic methane-oxidizing bacteria from northern peat wetlands. *Science* **282**:281–284.
- Dunfield, P. F., V. N. Khmelenina, N. E. Suzina, Y. A. Trotsenko, and S. N. Dedysh. 2003. *Methylocella silvestris* sp. nov., a novel methanotrophic bacterium isolated from an acidic forest Cambisol. *Int. J. Syst. Evol. Microbiol.* **53**:1231–1239.
- Dunfield, P. F., M. Tchawa Yimiga, S. N. Dedysh, U. Berger, W. Liesack, and J. Heyer. 2002. Isolation of a *Methylocystis* strain containing a novel *pmoA*-like gene. *FEMS Microbiol. Ecol.* **41**:17–26.
- Flanagan, J. L., L. Ventra, and A. S. Weiss. 1989. Rapid method for preparation and cleavage of bacterial DNA for pulsed-field gel electrophoresis. *Nucleic Acids Res.* **17**:814.
- Frishman, D., A. Mironov, H. W. Mewes, and M. Gelfand. 1998. Combining diverse evidence for gene recognition in completely sequenced bacterial genomes. *Nucleic Acids Res.* **26**:2941–2947.
- Gilbert, B. I., R. McDonald, R. Finch, G. P. Stafford, A. K. Nielsen, and J. C. Murrell. 2000. Molecular analysis of the *pmo* (particulate methane monooxygenase) operons from two type II methanotrophs. *Appl. Environ. Microbiol.* **66**:966–975.
- Hanson, R. S., and T. E. Hanson. 1996. Methanotrophic bacteria. *Microbiol. Rev.* **60**:439–471.
- Harley, C. B., and R. P. Reynolds. 1987. Analysis of *E. coli* promoter sequences. *Nucleic Acids Res.* **15**:2343–2361.
- Henikoff, S., and J. G. Henikoff. 1992. Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. USA* **89**:10915–10919.
- Heyer, J., V. F. Galchenko, and P. F. Dunfield. 2002. Molecular phylogeny of type II methane-oxidizing bacteria isolated from various environments. *Microbiology* **148**:2831–2846.
- Heyer, J., Y. R. Malashenko, U. Berger, and E. N. Budkova. 1984. Verbreitung methanotropher Bakterien. *Z. Allg. Mikrobiol.* **24**:725–744.
- Holmes, A. J., A. Costello, M. E. Lidstrom, and J. C. Murrell. 1995. Evidence that particulate methane monooxygenase and ammonia monooxygenase may be evolutionarily related. *FEMS Microbiol. Lett.* **132**:203–208.
- Klotz, M. G., J. Alzerreca, and J. M. Norton. 1997. A gene encoding a membrane protein exists upstream of the *amoA/amoB* genes in ammonia oxidizing bacteria: a third member of the *amo* operon? *FEMS Microbiol. Lett.* **150**:65–73.
- Miyaji, A., T. Kamachi, and I. Okura. 2002. Improvement of the purification method for retaining the activity of the particulate methane monooxygenase from *Methylosinus trichosporium* OB3b. *Biotechnol. Lett.* **24**:1883–1887.
- Murrell, J. C., B. Gilbert, and I. R. McDonald. 2000. Molecular biology and regulation of methane monooxygenase. *Arch. Microbiol.* **173**:325–332.
- Murrell, J. C., and S. Radajewski. 2000. Cultivation-independent techniques for studying methanotroph ecology. *Res. Microbiol.* **151**:807–814.
- Myers, E. W., and W. Miller. 1988. Optimal alignments in linear space. *Comput. Appl. Biosci.* **4**:11–17.
- Nguyen, H. H., S. J. Elliott, J. H. Yip, and S. I. Chan. 1998. The particulate methane monooxygenase from *Methylococcus capsulatus* (Bath) is a novel copper-containing three-subunit enzyme. Isolation and characterization. *J. Biol. Chem.* **273**:7957–7966.
- Nielsen, A. K., K. Gerdes, and J. C. Murrell. 1997. Copper-dependent reciprocal transcriptional regulation of methane monooxygenase genes in *Methylococcus capsulatus* and *Methylosinus trichosporium*. *Mol. Microbiol.* **25**:399–409.
- Osoegawa, K., P. Y. Woon, B. Zhao, E. Frengen, M. Tateno, J. J. Catanese, and P. J. de Jong. 1998. An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* **52**:1–8.
- Peterson, D. G., J. P. Tomkins, D. A. Frisch, R. A. Wing, and A. H. Paterson. 2000. *Journal of Agricultural Genomics*, vol. 5. Construction of plant bacterial artificial chromosome (BAC) libraries: an illustrated guide. CABI Publishing, Cambridge, Mass.
- Prior, S. D., and H. Dalton. 1985. The effect of copper ions on membrane content and methane monooxygenase activity in methanol-grown cells of *Methylococcus capsulatus* (Bath). *J. Gen. Microbiol.* **131**:155–163.
- Reeburgh, W. S., S. C. Whalen, and M. J. Alperin. 1993. The role of methylo-trophy in the global methane budget, p. 1–14. In J. C. Murrell and D. P. Kelly (ed.), *Microbial growth on C1 compounds*. Intercept Ltd., Andover, United Kingdom.
- Semrau, J. D., A. Chistoserdov, J. Lebron, A. Costello, J. Davagnino, E. Kenna, A. J. Holmes, R. Finch, J. C. Murrell, and M. E. Lidstrom. 1995. Particulate methane monooxygenase genes in methanotrophs. *J. Bacteriol.* **177**:3071–3079.
- Shine, J., and L. Dalgarno. 1974. The 3'-terminal sequence of *Escherichia*

- coli* 16S ribosomal RNA: complementarity to non-sense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. USA* **71**:1342–1346.
34. **Sobral, B. W., and A. G. Atherly.** 1989. A rapid and cost-effective method for preparing genomic DNA from gram-negative bacteria in agarose plugs for pulsed-field gel electrophoresis. *BioTechniques* **7**:938.
35. **Stolyar, S., A. M. Costello, T. L. Peebles, and M. E. Lidstrom.** 1999. Role of multiple gene copies in particulate methane monooxygenase activity in the methane-oxidizing bacterium *Methylococcus capsulatus* Bath. *Microbiology* **145**:1235–1244.
36. **Tchawa Yimga, M., P. F. Dunfield, P. Ricke, J. Heyer, and W. Liesack.** 2003. Wide distribution of a novel *pmoA*-like gene copy among type II methanotrophs, and its expression in *Methylocystis* strain SC2. *Appl. Environ. Microbiol.* **69**:5593–5602.
37. **Tukhvatullin, I., A., R. I. Gvozdev, and K. K. Anderson.** 2001. Structural and functional model of methane hydroxylase of membrane-bound methane monooxygenase from *Methylococcus capsulatus* (Bath). *Russ. Chem. Bull.* **50**:1867–1876.
38. **Whittenbury, R., K. C. Phillips, and J. F. Wilkinson.** 1970. Enrichment, isolation and some properties of methane-utilizing bacteria. *J. Gen. Microbiol.* **61**:205–218.
39. **Yang, Z., and J. P. Bielawski.** 2000. Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.* **15**:496–503.
40. **Zahn, J. A., and A. A. DiSpirito.** 1996. Membrane-associated methane monooxygenase from *Methylococcus capsulatus* (Bath). *J. Bacteriol.* **178**:1018–1029.

3.2 Vergleichende Untersuchung von Genom-Fragmenten eines nicht-kultivierten Vertreters des USC α sowie von *M. acidiphila*

Übersicht :

Bis heute konnten keine MB isoliert werden, die zu Wachstum unter atmosphärischen Methan-Konzentrationen befähigt wären. Kultivierungsunabhängige Untersuchungen der methanotrophen Gilden als Methan-Senken charakterisierter Upland-Böden ergaben die Abundanz bislang nicht kultivierter MB des USC α . Hochmolekulare DNA der mikrobiellen Population eines sauren Waldbodens wurde extrahiert und für die Konstruktion einer Metagenom-Bibliothek genutzt. Mittels eines *pmoA*-spezifischen PCR-Assays wurden die 250.000 Klone der Umwelt-Bibliothek nach Genom-Fragmenten der potentiell hoch-affinen MB des USC α untersucht. Das Screening resultierte in der Identifikation zweier Klone, deren Inserts eindeutig dem USC α zugeordnet werden konnten. Die vergleichende Analyse eines 42 kb großen Contigs ermöglichte die eindeutige phylogenetische Zuordnung des USC α zu den Rhizobiales (*Alphaproteobacteria*) mit hohen Ähnlichkeiten zu den Genomen der *Bradyrhizobiaceae*. Als Referenz wurde eine Genom-Bibliothek des acidophilen MB *M. acidiphila* konstruiert und ein 100 kb großes genomisches Fragment sequenziert. Die detaillierte phylogenetische Analyse von auf beiden analysierten Genom-Fragmenten nachgewiesenen Genen ergab eine nahe Verwandtschaft von USC α und *M. acidiphila*. Die Analyse der für die pMMO kodierenden Gene ergab für beide analysierten Genom-Fragmente die Anordnung in einem Operon in der Reihenfolge *pmoCAB*. Es konnten auf molekularer Ebene keine Hinweise für eine erhöhte Affinität der pMMO des USC α zu Methan gefunden werden. Allerdings konnte stromabwärts des *pmo*-Operons ein Gen identifiziert werden, das möglicherweise ein viertes Gen des *pmo*-Operons darstellt. Darauf deuten hohe Sequenz-Ähnlichkeiten zu Genen hin, die sowohl bei *M. trichosporium* Stamm OB3b als auch bei den homologen AMO-Operons nitrifizierender Bakterien ebenfalls stromabwärts des *pmoB/amoB* Gens lokalisiert sind.

Anmerkung :

Die Shotgun-basierte Sequenzierung der Inserts ausgewählter Klone wurde im Rahmen des Kompetenznetzwerks GenoMik am MPI für molekulare Genetik, Berlin durch Dr. M. Kube durchgeführt. Auch die automatisierte Primär-Annotation erfolgte in Berlin.

USC α , a group of as-yet uncultured methanotrophs:**First insights into the genomic organization substantiates an evolutionary affiliation to the alphaproteobacterial *Methylocapsa acidiphila***

Peter Ricke, Michael Kube, Christoph Erkel, Richard Reinhard, and Werner Liesack

Submission to Applied and Environmental Microbiology intended

Abstract

Methanotrophic bacteria (MB) of the "upland soil cluster α " (USC α) are assumed to be responsible for methane oxidation at atmospheric methane concentrations but have as yet escaped all cultivation attempts. A metagenomic fosmide library consisting of 250.000 clones was constructed from total DNA of an acidic forest soil previously characterized as a sink for atmospheric methane. Using *pmoA*-targeted PCR assays, two clones were identified to carry genomic fragments of USC α representatives. Shotgun-based sequencing revealed that both inserts belonged to the same strain or two highly related USC α genotypes. Thus sequences were assembled resulting in a 42-kb contig. The next cultured relative of USC α is, based on partial *pmoA* phylogenies, the acidophilic *M. acidiphila*. Consequently, a 100-kb fragment of the *M. acidiphila* genome containing the *pmo* operon was sequenced. Annotation and comparative analysis of both genomic fragments against publicly available completely sequenced genomes revealed highest similarities to members of the *Bradyrhizobiaceae*. Phylogenetic analyses of four genes identified on both fragments corroborated a close phylogenetic relation between *M. acidiphila* and USC α . The affiliation of USC α to the *Alphaproteobacteria* was also confirmed by a SOM-based analysis of di-, tri- and tetra-nucleotide patterns of the USC α genomic fragment. A detailed comparative analysis of the *pmo* operons of USC α and *M. acidiphila* showed that both the *pmo* operon structure and the predicted secondary structure of pMMO are highly conserved among all MB. No particularities were identified at the DNA or amino acid level that might have been interpretable as indicators for the putative capability of USC α to oxidize methane with high affinity. Together with the recently reported *pmoCAB1* and *pmoCAB2* of *Methylocystis* strain SC2, this study doubled the number of completely analysed pMMO operons. It enabled a phylogenetic analysis of concatenated *pmoCAB* sequences rather than only of partial *pmoA*, thereby roughly sextupling the amount of underlying genetic information.

Introduction

Aerobic methanotrophic Bacteria (MB) are able to use methane (CH_4) as sole source of carbon and energy. Based on phylogenetic, physiological, morphological, and biochemical characteristics, methanotrophs have been divided into three phylogenetically and physiologically distinct groups, type X, type I and type II MB, forming coherent clusters within the Gammaproteobacteria (type X and type I) and Alphaproteobacteria (type II) (Hanson and Hanson, 1996). The gammaproteobacterial type I MB comprises the genera *Methylomonas*, *Methylocaldum*, *Methylomicrobium*, *Methylobacter*, *Methylosarcina*, and *Methylosphaera*. While the type X MB *Methylococcus* also belong to the Gammaproteobacteria the alphaproteobacterial type II MB consists of the genera *Methylocystis* and *Methylosinus*. With the advent of novel MB isolates and environmental studies, this classification has been challenged. Novel isolates were obtained, that could not be assigned into the type-system. For example, the recently described acidophilic MB that are represented by two genera, *Methylocella* and *Methylocapsa*, possess several unique morphological and physiological characteristics, and, based on 16S rRNA phylogeny, *Methylocella* and *Methylocapsa* are evolutionarily more closely related to acidophilic heterotrophic bacteria of the genus *Beijerinckia* than to the *Methylosinus/Methylocystis* group.

The first step in CH_4 oxidation, the conversion of methane to methanol, is carried out by a methane monooxygenase (MMO). This enzyme exists in two forms, a particulate, membrane-associated form (pMMO) and a soluble form (sMMO). The two forms of the enzyme differ in structure, in kinetic properties, and in the range of substrates which are utilized. Only a restricted number of MB species possess sMMO, while almost all MB possess pMMO. The only MB lacking pMMO are members of the genus *Methylocella*. In MB that harbor both forms of MMO, sMMO is synthesized under copper-deficient conditions, while in the presence of even a minuscule amount of available Cu(II) (0.85 to 1.0 $\mu\text{mol/g}$ [dry weight] of cells) only pMMO is synthesized.

The pMMO is composed of two subunits, a hydroxylase and a NADH-oxidoreduktase. The hydroxylase subunit has been demonstrated to be encoded by three consecutive open reading frames (*pmoC*, *pmoA*, and *pmoB*) in both type X MB and type II MB. The *pmo* operon is homologous to the *amoCAB* operon that encodes ammonium monooxygenase (AMO), the key enzyme of autotrophic nitrifiers. AMO catalyzes not only the oxidation of ammonia but also that of CH_4 , albeit at a much lower activity.

PmoA and AmoA are presumed to contain the active sites of the respective enzymes because they have been shown to be specifically labeled by [^{14}C] acetylene, a suicide substrate for MMO and AMO. *pmoA* has been used in numerous studies on the diversity and distribution of MOB, as the phylogeny of *pmoA* sequences shows a strong congruity with that of 16S rDNA sequences (Murrell et al., 1998). It is an ideal marker for environmental studies because it enables the functional detection of novel, as-yet uncultured MB.

The only biological sink for CH_4 is oxidation in soil. Atmospheric CH_4 is consumed in forest, agricultural, and other upland soils. CH_4 consumption in these soils is caused by MB which are characterized by a high apparent affinity for methane of 10–200 nM (reviewed by (Dunfield et al., 1999). In forest soils that are sinks for atmospheric CH_4 , novel *pmoA* sequence types (referred to as "upland soil cluster α " or "USC α ") have been described frequently. More recently, an additional novel *pmoA* lineage was detected in upland soils. Since it grouped distantly related to type I MB in *pmoA*-based phylogenies, it was designated "upland soil cluster γ " (USC γ). USC γ was detected in soils with pH values greater than 6.0, whereas *pmoA* of USC α has been retrieved mainly from acidic soils. No pure cultures are known for these putative atmospheric methane oxidizers.

Phylogenies constructed for partial *pmoA* sequences group USC α sequences within *pmoA* sequences of alphaproteobacterial origin. The inferred amino acid sequences of USC α exhibited closest relatedness to PmoA of *Methylocapsa acidiphila*. In phylogenetic trees USC α and *M. acidiphila* partial PmoA sequences formed a common branch, clustering distinct from PmoA sequences of conventional type II MB of the *Methylosinus/Methylocystis* group (Dedysh et al., 2001). Thus, based on *pmoA* phylogenies *M. acidiphila* has been regarded as the closest cultivated relative of USC α . However, the apparent affinity for methane exhibited by *Methylocapsa acidiphila* B2 was 1–2 μM , which is similar to values measured in other methanotrophic bacteria.

Since 16S rDNA-based phylogenetic information is still unavailable for USC α , its true evolutionary origin remained to be elucidated. We used a metagenomic approach to obtain a first insight into the genome of USC α representatives. Besides final elucidation of the evolutionary origin, the retrieval of USC α genomic fragments also aimed at assessing whether the pMMO of USC α representatives exhibits any unusual feature(s) that correspond(s) to the assumed capability of USC α to consume atmospheric methane. The *pmo* operon sequence and flanking genomic regions of *M. acidiphila* were analysed for comparison.

Materials and Methods

Soil samples and bacterial strains. Soil (100 g of Braunerde (Cambisol), pH = 4.3) was collected from a forest located near Marburg, Germany in October 2003 und April 2004. After removal of rotten organic material, sampling from a soil depth of a 3 to 15 cm (A_h) was performed at a beech grove (*Fagus*, *Quercus*). The soil was stored at -20°C . Characteristics regarding the MOB population of of the sampled soil have been reported by Knief et al. (Knief et al., 2003). The conditions used for growth of *M. acidiphila* B2 were adapted from Heyer et al. (10, 11). Cells were grown in liquid culture for 3 to 4 weeks at 30°C under a headspace of 10 % (vol/vol) CH_4 , 3 % CO_2 , and 87 % air.

Extraction of high-molecular-weight (HMW) DNA from forest soil. To each 4 g of forest soil, 3 ml sodium phosphate buffer (pH = 8.0) and 250 μl of 100 mg/ml lysozyme stock solution were added. After incubation for 1 h at 37°C , 60 μl of Proteinase K (20 mg/ml, Qiagen, Hilden, Germany) were added and incubation was continued for 1 h. Samples were frozen in liquid nitrogen and stored for 0.5 h at -80°C . After thawing, 400 μl 20% (w/v) SDS were added, the samples were mixed gently and incubated at 65°C for 45 min. Samples were centrifuged at 3500 rpm for 5 min in a falcon-tube centrifuge, supernatants were centrifuged for 1 min at 10,000 rpm in a microcentrifuge. Supernatants were combined with 0.45 % (v/v) 6 M potassium acetate, incubated for 5 min at RT and proteine and cell residuals were precipitated at 13,000 rpm for 5 min. To reduce the humic acid content, DNA extracts were further purified by use of the FastDNA[®] Spin[®] Kit for Soil (Q-BIOgene, Heidelberg, Germany) according to the manufacturers instructions with slide modifications: four additional washing steps with 500 μl of 5,5 M guanidin-thiocyanat were performed after the DNA was loaded on the spin columns. Purified DNA was end-repaired and size fractionated via PFGE using a protocol described previously (17-19). Fragment sizes between 35 and 50 kb were eluted immediately from the agarose gel by agarase digestion (β -agarase I; NEB, Beverly, USA). DNA was concentrated by vacuum centrifugation to a final concentration of approx. 20 ng/ μl .

Construction of a metagenomic fosmid library. Fosmid clones were produced using the CopyControlTM Fosmid Library Production Kit (Epicentre, Madison, USA) as described by the manufacturer with slide modifications as follows: i) a second end-repair reaction was performed after size selection, ii) ligation was extended over night at 4°C , and iii) each

phage packing reaction was extended to 3 h. The average insert size of the resulting clones was estimated by *Sbf*I restriction and PFGE analysis of the vector/insert constructs of 20 randomly chosen clones as described by Peterson et al. (18).

Construction of genomic BAC library from *M. acidiphila* B2. HMW DNA is a prerequisite for construction of a genomic BAC library. In order to prevent shearing of the DNA, whole cells of *M. acidiphila* were embedded in agarose at a final cell density of $5 \times 10^9 \text{ ml}^{-1}$. Cells in the agarose blocks were lysed in SDS-based lysis buffer (Dedysh et al., 1998), followed by partial digestion of HMW DNA with *Hind*III endonuclease (New England Biolabs, Frankfurt, Germany) according to Peterson et al. (2000). Size fractionation was performed by PFGE. HMW DNA of approximately 50 to 150 kb in size was eluted by agarase digestion as described above and ligated to 100 ng of pBeloBAC11 (Epicentre Technologies, Madison, Wisconsin, USA). Best cloning efficiencies were obtained at a molar ratio of insert to vector DNA of approximately 1:10. Ligation was followed by transformation of the desalted DNA into *E. coli* cells (ElectroMax DH10B competent cells; Invitrogen Life Technologies; Carlsbad, USA) as described elsewhere (Ricke et al., 2004). Single clones were picked and transferred into 96-well microtiter plates. The average insert size was estimated as described above.

PCR-based screening. The screening for *M. acidiphila* BAC clones carrying *pmo* genes was performed by combining individual clones to mixtures of twelve (“row-pools”) and 96 (“plate-pools”) clones, followed by cell lysis and PCR amplification of *pmoA* as described previously (19). Primers 189f (GGN GAC TGG GAC TTC TGG) and 662b (GAA SGC NGA GAA GAA SGC) were used with a PCR protocol described by Holmes et al. (Holmes et al., 1995). To proof the correct identification of test-positive clones, amplicons of the expected size were purified and sequenced by application of the Big Dye technology (PE Applied Biosystems, Weiterstadt, Germany).

To be able to rapidly process a large number of fosmid clones, a fast high throughput method was established for screening of the forest soil metagenomic library. The colonies grown on the individual plates were directly replicated by using the Lederberg technique. Colonies were washed from the replicated agar plates with steril water and thoroughly mixed by vortexing for 1 min. Cell aliquots (100 μl) were lysed by boiling and 1 μl each was used as template for PCR-based screening. Residual cells were frozen with 1 volume of glycerol buffer (65% [v/v] Glycerol, 1,2 mM MgSO_4 , pH 8,0) in liquid nitrogen and

stored at -80° C. Screening for *pmoA*-carrying inserts was performed by applications of the above described PCR assay and a USC α specific PCR assay reported by Kolb et al. (Kolb et al., 2003).

DNA sequencing and computational analysis. The inserts of two *pmoA*-positive clones of *M. acidiphila* and two test-positive environmental clones were characterized by a standard procedure, including shotgun-cloning, sequencing (Phred 30), assemblage and annotation. Annotation was carried out with the Software packages Orpheus (gene-prediction, (<http://pedant.gsf.de/orpheus/>), HTGA (high throughput gene annotation, (Rabus et al., 2002) and Artemis (final manual annotation, <http://www.sanger.ac.uk/Software/Artemis/>). Homology searches were carried out with standard software tools, including the BLAST algorithm (Altschul et al., 1990) against non-redundant DNA database as well as against a non-redundant protein database compiled for this study from SWISSPROT, TrEMBL and PIR. Highly conserved amino acid sequence motifs were identified by comparison against the InterPro database of protein families, domains and functional sites (<http://www.ebi.ac.uk/interpro/>). The locations of putative transmembrane-spanning regions of PmoCAB were computed using the programs toppred (<http://bioweb.pasteur.fr/seqanal/interfaces/toppred.html>) and TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>). Similarity and identity values were calculated on the basis of global pairwise alignments (Myers and Miller, 1988) using the stretcher program (<http://bioweb.pasteur.fr/seqanal/interfaces/stretcher.html>) and appropriate BLOSSUM (Henikov and Henikoff, 1992) substitution matrices.

All phylogenetic trees were calculated within the ARB program package (developed by O. Strunk and W. Ludwig of the Technical University of Munich [<http://www.arb-home.de>]). Comparative analyses of the *pmo* genes and derived amino acid sequences were based on manually corrected multiple alignments that were computed with the ClustalW algorithm. For analysis of the complete *pmoCAB* operons, sequences of PmoCAB and AmoCAB of individual organisms were concatenated while omitting intergenic, noncoding regions. For analysis of other protein sequences, multiple alignments of protein families were obtained from the Pfam UK database. Maximum likelihood (ProtML and TreePuzzle) and neighbor-joining methods were applied using appropriate evolutionary models. The resulting phylogenies were combined manually into a consensus tree while ambivalent branches were shown as a multifurcation.

SOM based Mapping. Analysis of genomic signatures was performed by application of the XanaMap program (<http://www.xanagene.com>). Genomic fragments of USC α and *M.acidiphila* were divided into 5 kb subfragments and compared against signature maps computed for all completely sequenced genomes available provided by the XanaMap program. The unsupervised neural network algorithm self-organizing map (SOM) is discussed in detail below.

Nucleotide accession numbers. The annotated 100-kb fragment of *M. acidiphila* B2 has been deposited in the EMBL, DDBJ and GenBank databases under the accession number xxxxxxxx. The genomic fragment of the USC α representative was assigned the accession number xxxxxxxx.

Results

Construction and screening of large insert libraries (LILs). The major aims of this study were i) to obtain first insights into the genome of USC α representative(s), ii) to determine the evolutionary origin of these as-yet uncultured methanotrophs, and iii) to perform a detailed analysis of the *pmo* operon. The rationale behind focusing on the third aspect was to check whether the pMMO of USC α exhibits at the molecular level any unusual features (characteristics) that might be linked to the putative capability of this methanotroph group to oxidize methane at atmospheric concentrations. Since no pure-culture representatives and even no enrichments of USC α are known, a metagenomic approach was performed using the fosmid technology for construction of the environmental clone library. For comparison of sequence information, a large insert genomic library was constructed also for *M. acidiphila*, the next cultured relative of USC α based on *pmoA*-phylogenies.

Cloning resulted in two LILs: i) a forest soil metagenomic fosmid library that consisted of approximately 250.000 clones with an average insert size of 40 kb and ii) a *M. acidiphila* library that consisted of 1054 BAC clones with an average insert size of 70 to 80 kb. PCR-based screening led to the identification of two environmental clones carrying an USC α *pmoA* gene. Shotgun-based sequencing revealed that both inserts belonged to the same strain or two highly related USC α genotypes, as the obtained sequences were identical over the complete overlapping region of 30.000 bp. Assemblage of the two genomic

fragments resulted in a 42-kb contig. Eight clones of the *M. acidiphila* genomic BAC library were identified to carry the *pmoA* gene. A restriction analysis of the cloned inserts (data not shown) resulted in the identification of two genomic fragments with a minimal overlap for shotgun-based sequencing. Finally, a 101-kb contig of non-redundant sequence information was obtained.

Annotation of an USC α 42-kb genomic fragment. Of 71 open reading frames (ORFs) identified on the USC α genomic fragment by the ORF prediction program ORPHEUS (6), 47 ORFs were predicted to encode proteins. Manual analysis in combination with semi-automatical annotation by the HTGA platform enabled the assignment of 27 ORFs to genes with known functions, while 20 ORFs were referred to as hypothetical or conserved hypothetical. The designation "conserved hypothetical" (5 ORFs) refers to ORFs with high similarity to predicted genes with an as-yet unknown function. The coding density of the USC α fragment was calculated to be 85.0 %. The ORF length averages 767 bp and the mean mol% G+C content is 59.8 %.

Thirty-two ORFs had at least one homolog sequence in public databases. Four of them are arranged in a gene cluster containing the *pmoCAB* operon plus an additional gene of unknown function. Genes of this cluster are arranged in the same organisation as previously reported for other MB and autotrophic nitrifiers (see below). Based on BLAST analyses, 26 of the remaining 28 ORFs showed highest similarities to members of the *Alphaproteobacteria*. Fourteen of them exhibited highest similarities to genes from *B. japonicum*.

A different subset of 14 ORFs was arranged in four gene clusters with high structural similarity to genomic regions of *B. japonicum* (Fig.1). Three of the four clusters were also found in *R. palustris*, which is also a member of the *Bradyrhizobiaceae*. The genome of *B. japonicum* was the only completely sequenced genome in which homologs of all four USC α gene clusters were present. No USC α gene cluster was found to be conserved in any other genome, if it was not also present in the *B. japonicum* genome. However, for each of the four USC α gene clusters at least partially similar genomic regions could be identified in more than one publicly available genome. In contrast to homologous *B. japonicum* clusters, which could be identified easily based on BLAST similarity searches, most of the homologous gene clusters identified in other bacteria were found only by analysis of manually curated databases of orthologous genes (Kegg-KO, NCBI-COG). An identification based on simple automated similarity comparison methods was not

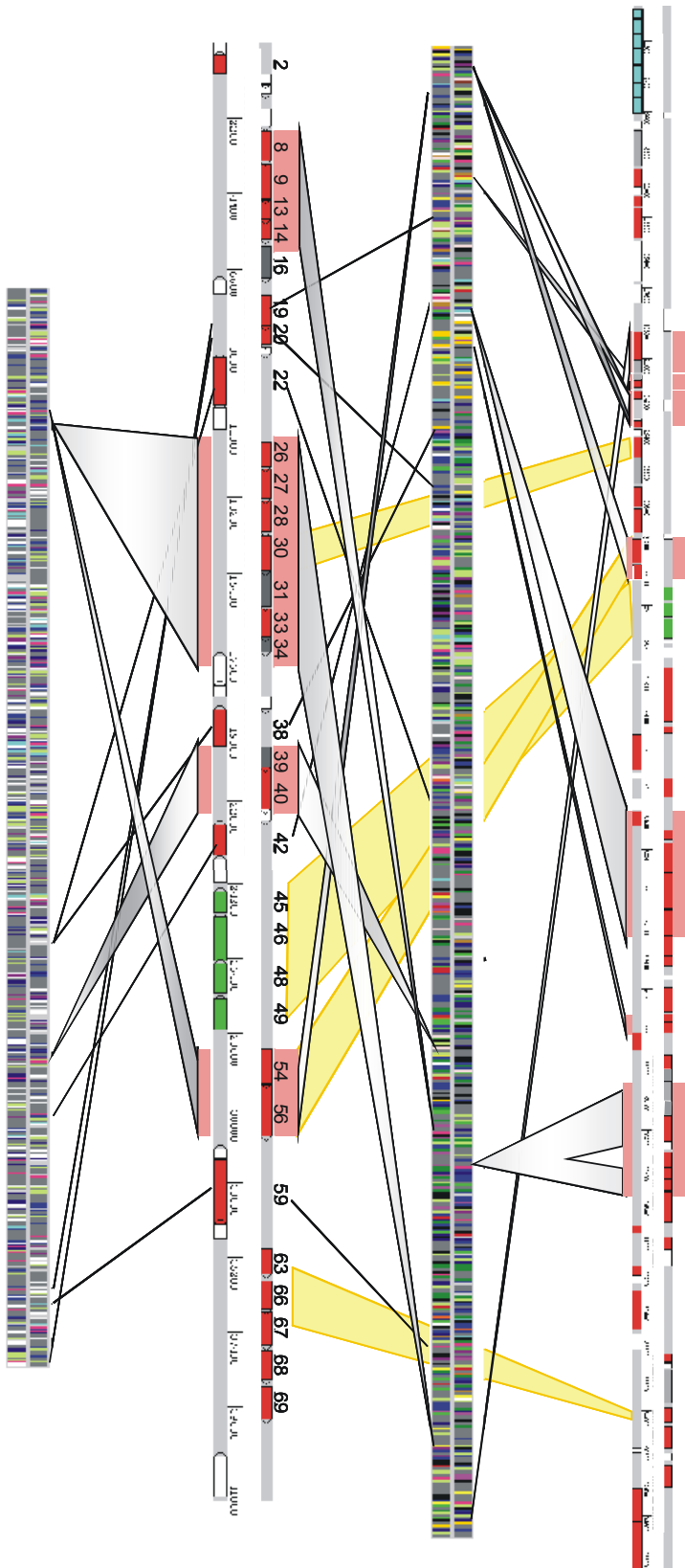


Figure 1. Schematic overview of gene arrangements for i) the 42 -kb fragment of USC α , ii) the 100 -kb fragment of *M. acidiphila* strain B2, and iii) the complete genomes of *B. japonicum* and *R. palustris*. Corresponding regions are linked by shaded interconnections. T he numbering of ORFs corresponds to that used in the Text. Colour code: green: *pmoCAB* and the associated ORF45 (OrtD, Gilbert et al, 2001); blue: g enes encoding enzymes involved in H4MPT /MFR-mediated C1 metabolism and homologous to the gene cluster described by Vorholt et al (1999) for *M. extorquens*; red: other proteins with assigned functions; white: hypothetical proteins; grey: conserved hypothetical proteins.

possible. As the conservation of structural genomic features is a good indicator for a common evolutionary origin of the organisms (or at least of the genomic regions) under comparison, the four highly conserved USC α gene clusters were characterised in detail.

"Cluster 1" consists of four genes that encode the putative general secretory pathway proteins GspN, GspM, GspL and GspK. In *B. japonicum* and four *Gammaproteobacteria* of the genera *Xylella* and *Xanthomonas*, these genes are part of a conserved cluster of up to 12 genes coding for secretory pathway proteins. In contrast, these genes are missing in most completely sequenced alphaproteobacterial genomes, e.g. the genome of *R. palustris*. The order of the genes *gspN*, *gspM*, *gspL* and *gspK* resembles even more closely those of the aforementioned *Gammaproteobacteria* than the homologous gene cluster in the *B. japonicum* genome, as the latter possesses a gene insertion between *gspN* and *gspM*.

"Cluster 2" consists of seven genes. The ORFs 26, 27 and 28 encode proteins involved in zinc/manganese ABC transporter system: the ATP binding protein (ORF26), the ABC transporter permease (ORF27) and the periplasmic substrate-binding protein (ORF28). The organisation of this gene cluster is highly conserved among bacteria and is partially conserved even in some archaea. On the UCS α fragment, two genes (ORF30 and ORF31) were identified directly downstream of this cluster, which are located in a homologous position also in some alphaproteobacteria, betaproteobacteria and cyanobacteria. An additional gene encoding acetoacetate decarboxylase (ORF33) and a paralogous gene copy of ORF 31 (ORF34) complete the gene cluster. The acetoacetate decarboxylase gene is found only in *B. japonicum* genome upstream of the aforementioned genes, so that the organisation of the complete cluster resembles strongly those found in UCS α .

"Cluster 3" consists of two genes. ORF39 was annotated as conserved hypothetical gene and, so far, identified only in *Alphaproteobacteria*. The second gene (ORF40) codes for salicyhydroxylase. The gene organization of cluster 3 is found only in *Alphaproteobacteria*, including *B. japonicum*, *R. palustris*, and *Brucella* spp.

"Cluster 4" is localized directly upstream of the *pmoCAB* gene cluster and is formed by two genes, ORFs 54 and 56. Genes with similarity to ORF54 were found by Blast analysis only in *B. japonicum* and *R. palustris*. Both genes were annotated as hypothetical proteins. Analysis of orthologous genes and comparison against the InterPro database enabled their annotation as members of the "Staphylococcus nuclease subtype" protein family, which is widespread among bacteria, archaea and eukaryotes. ORF56 was identified

as a zinc-metallo protease. Besides *B. japonicum* and *R. palustris*, only *P. aeruginosa* was found to carry homologous genes in the same gene order. However, the similarity between the corresponding gene of *P. aeruginosa* and ORF54 is very low. Interestingly, an homologous gene cluster was identified also on the *M. acidiphila* genomic fragment directly upstream of to the *pmoCAB* operon. This agrees well with the localization of cluster 4 on the USC α genomic fragment.

In contrast to the overall high degree of similarity between the USC α fragment and genomes of the *Bradyrhizobiaceae*, a single ORF (ORF16) was identified whose derived amino acid sequence exhibited no significant similarities to any alphaproteobacterial sequence but instead was highly similar to a hypothetical protein of *Nitrosomonas europaea*. Also, the independent SOM-based analysis (see below) of the USC α fragment predicted a betaproteobacterial origin for the genomic region containing ORF16.

Annotation of a *M. acidiphila* 101-kb genomic fragment. In total, 149 ORFs with sizes above 200 bp were identified on the genomic fragment of *M. acidiphila*. Detailed annotation identified 90 ORFs that were predicted to encode proteins. Fifty-eight ORFs could be assigned a putative function. For example, genes involved in DNA replication and recombination (chromosomal replication initiator protein DnaA, RuvC and RuvA) and genes encoding proteins of the ABC transporter family were identified. Thirty-two ORFs were referred to as hypothetical (24 ORFs) or conserved hypothetical genes (8 ORFs). The coding density of the *M. acidiphila* genomic fragment was calculated to be 86.9 % with an average ORF length of 989 bp. The G+C content is 62.0 %.

As for the USC α genomic fragment, BLAST analysis of predicted ORFs resulted in a high number of genes with closest similarities to those of *B. japonicum*. Altogether, 48 derived protein sequences exhibited clear similarities to genes of *B. japonicum*. Twenty-two genes were found to be organized in six gene clusters with similar structural organisations as exhibited by their homologs in the *B. japonicum* genome (Fig. 1). The largest cluster consists of the ORFs 78, 82, and 84 to 87. ORFs 84 to 87 encode for proteins that belong to the ABC transporter family. Of special interest is ORF78. It encodes a putative prephenate dehydratase and is a homolog to ORF42 of the USC α fragment. As mentioned above, two genes located directly upstream of the *pmoCAB* operon (ORFs 53 and 55) were also found on the USC α fragment. Moreover, ORF133 and ORF46 exhibited high similarities to ORFs identified on the USC α fragment. ORF133 was predicted to encode an outer membrane protein and ORF46 was annotated as *cowB*, encoding a

cobalamin synthesis protein. The occurrence of these five genes on the genomic fragments of both USC α and *M. acidiphila* enabled a more detailed analysis of the phylogenetic relationship between these two representatives of unusual MOB (see below).

Genes encoding enzymes involved in H₄MPT /MFR-mediated C₁ metabolism. In addition to the aforementioned genes, we identified archaeal-like genes that encode enzymes involved in H₄MPT/MFR-mediated C₁ transfer reactions. These enzymes, which are otherwise specific to methanogenic and sulfate-reducing archaea, have previously been shown to be involved in the formaldehyde oxidation to CO₂ in the methylotrophic *Methylobacterium extorquens* AM1 (Chistoserdova et al., 1998). Thus, in correspondence to various other methylotrophic and methanotrophic bacteria, *M. acidiphila* seems to possess a H₄MPT-dependent formaldehyde oxidation pathway. However, the comparison with the corresponding genomic region of *M. extorquens* suggests that the gene clusters predicted to code for enzymes involved in H₄MPT/MFR-mediated C₁ transfer reactions are differently organised in both organisms. In contrast to the genomes of *M. extorquens* or *M. capsulatus* (Bath), genes encoding the formyltransferase/hydrolase complex (Fhc A-D) are not flanking ORFs 1-12 but instead are missing on the *M. acidiphila* genomic fragment analysed here.

Phylogenetic analysis of genes present in the genomes of both USC α representative and *M. acidiphila*. Besides *pmoCAB* (see below), five additional genes were located on the genomic fragments of both USC α and *M. acidiphila*. Four of them were suitable for detailed phylogenetic analyses of the relationship between *M. acidiphila* and USC α , while too many paralogous genes were identified among the *Bradyrhizobiaceae* and the USC α fragment for the fifth gene (encodes a putative outer membrane protein). Regardless of the set of homologs used for comparison, all analysed datasets contained, in addition to the respective genes of USC α and *M. acidiphila*, homologous gene sequences of *B. japonicum* and *R. palustris*. The comparative genomic analyses of either USC α or *M. acidiphila* against the completely sequenced genomes available in public databases had resulted for both genomic fragments highest similarities to *B. japonicum* and *R. palustris*. Comparative analysis of the four different gene types that were present in the genome of all four organisms allowed us to resolve the phylogenetic relationships within this quartet group.

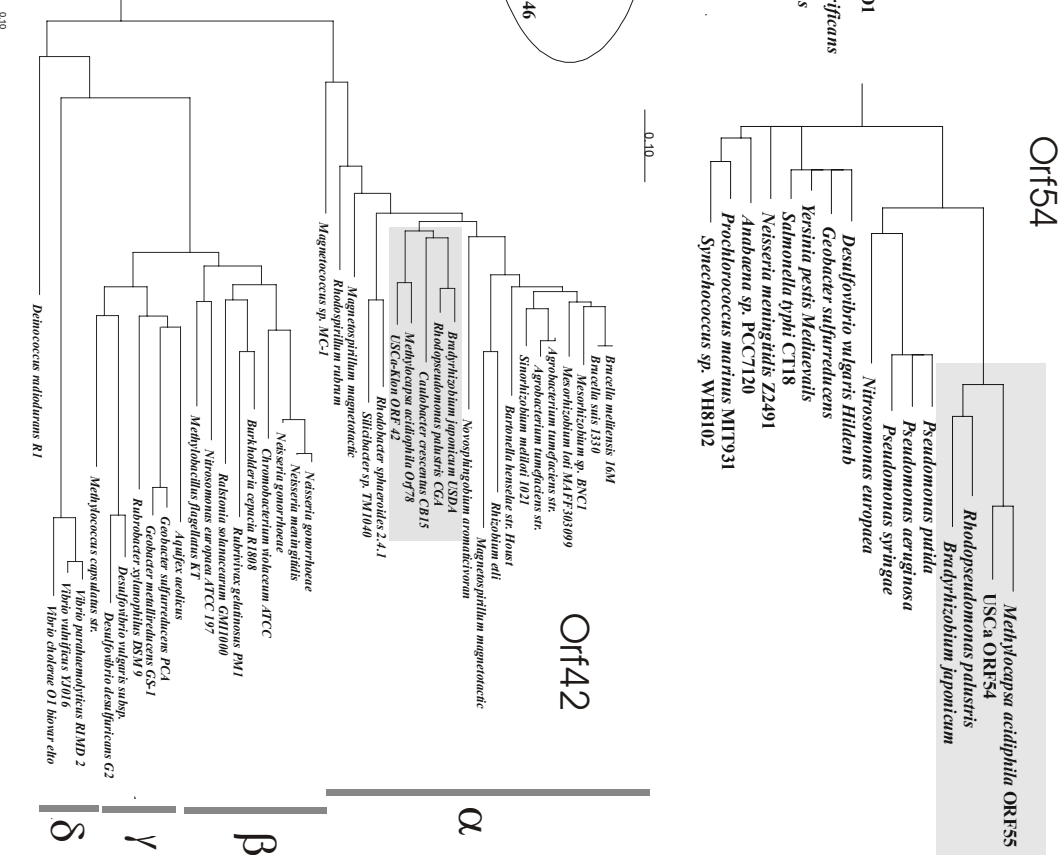


Figure 2. Phylogenetic analyses of polypeptides inferred from genes that were present on the genomic fragments of both USC α and *M. acidiphila*. The grey-shaded boxes indicate the identical branching pattern observed between the different gene phylogenies for USC α , *M. acidiphila*, *B. japonicum* and *R. palustris*. Dendrograms are manually multifurcated consensus trees of phylogenies constructed with distance-based as well as position-specific algorithms (Neighbour joining, TreePuzzle, ProtML) implemented in the ARB package (Ludwig et al. 2004).

Alignments of reference sequences included in the respective analysis could be obtained from the Pfam protein families database of alignments and HMMs (<http://www.sanger.ac.uk/Software/Pfam/>), or the Clusters of Orthologous Groups Database (COG, <http://www.ncbi.nlm.nih.gov/COG/>). Phylogenies of genes encoding the staphylococcal-like nuclease (PF00565), a putative zinc metallo protease (sequences retrieved from Kegg-KO), as well as a putative prephenate dehydratase (PF00800) always grouped the genes of USC α and *M. acidiphila* closely together within lineages formed by various alphaproteobacteria genes. The USC α /*M. acidiphila* branch always clustered together with a "sister lineage" formed by *B. japonicum* and *R. palustris* with the exception that in the tree reconstructed for the prephenate dehydratase sequences an additional organism (*C. crescentus* CB15) was assigned to this group (Fig. 2). The only phylogeny in which USC α and *M. acidiphila* sequences were clearly separated was that of putative CowB proteins. It could be clearly shown, that the compared genes of USC α and *M. acidiphila* are closely related to two different paralogous genes of *Mesorhizobium loti* and therefore were not applicable for phylogenetic inference.

Signature-based phylogenetic mapping of the USC α genomic fragment. Based on calculated pylogenies, similarities of the derived amino acid sequences as well as the genome structures, it could be shown that the USC α representative studied here belongs phylogenetically to the *Alphaproteobacteria*. For verification of this assumption, the genomic fragment of USC α was analysed with the XanaMap Software. This program performs a highly reliable phylogenetic assignment of genomic fragments to phylogenetic groups or even species based on a statistical analysis of the relative abundances of di-, tri- and tetranucleotides in a given sequence (Abe et al., 2002; Abe et al., 2003; Kanaya et al., 2001). The basis of XanaMap is an unsupervised neural network algorithm, self-organizing map (SOM), which can cluster complex, n-dimensional data (e.g. the frequencies of the analysed oligonucleotides in different fully sequenced genomes) and display them as two-dimensional maps. The individual pads of the resulting map thereby represent weight-vectors (data-points within the high-dimensional data space). The distance between two weight vectors reflects (in a non-linear matter) the differences in genomic signatures. This method had been shown to be adequate for classification of DNA sequences into subgroups that correspond to biological categories (taxa). For comparison with the USC α fragment, the genomic fragment of *M. acidiphila* and also non-redundant genomic sequences of *Methylosinus* spp. and *Methylocystis* spp. available in public databases (GenBank, EMBL, and DDBJ) were included in the analysis.

The signature analysis clearly placed all 38 5-kb USC α subfragments within an area characteristic for *Alphaproteobacteria* (Fig. 3) with only one exception placed within a map area dominated by *Betaproteobacteria*. This fragment contains ORF16 whose derived amino acid sequence exhibited high similarities only to a single GenBank database entry, a putative protein of the betaproteobacterial *Nitrosomonas europaea* (accession number: Q82s51). The ORFs flanking ORF16 showed relatedness to the *Alphaproteobacteria*, which is reflected by the placement of this 5-kb subfragment at the boundary between the alpha- and betaproteobacterial map area. On the species level, the USC α 5-kb subfragments were assigned to genomic fragments of *B. japonicum*, *S. meliloti* or *A. tumefaciensis*.

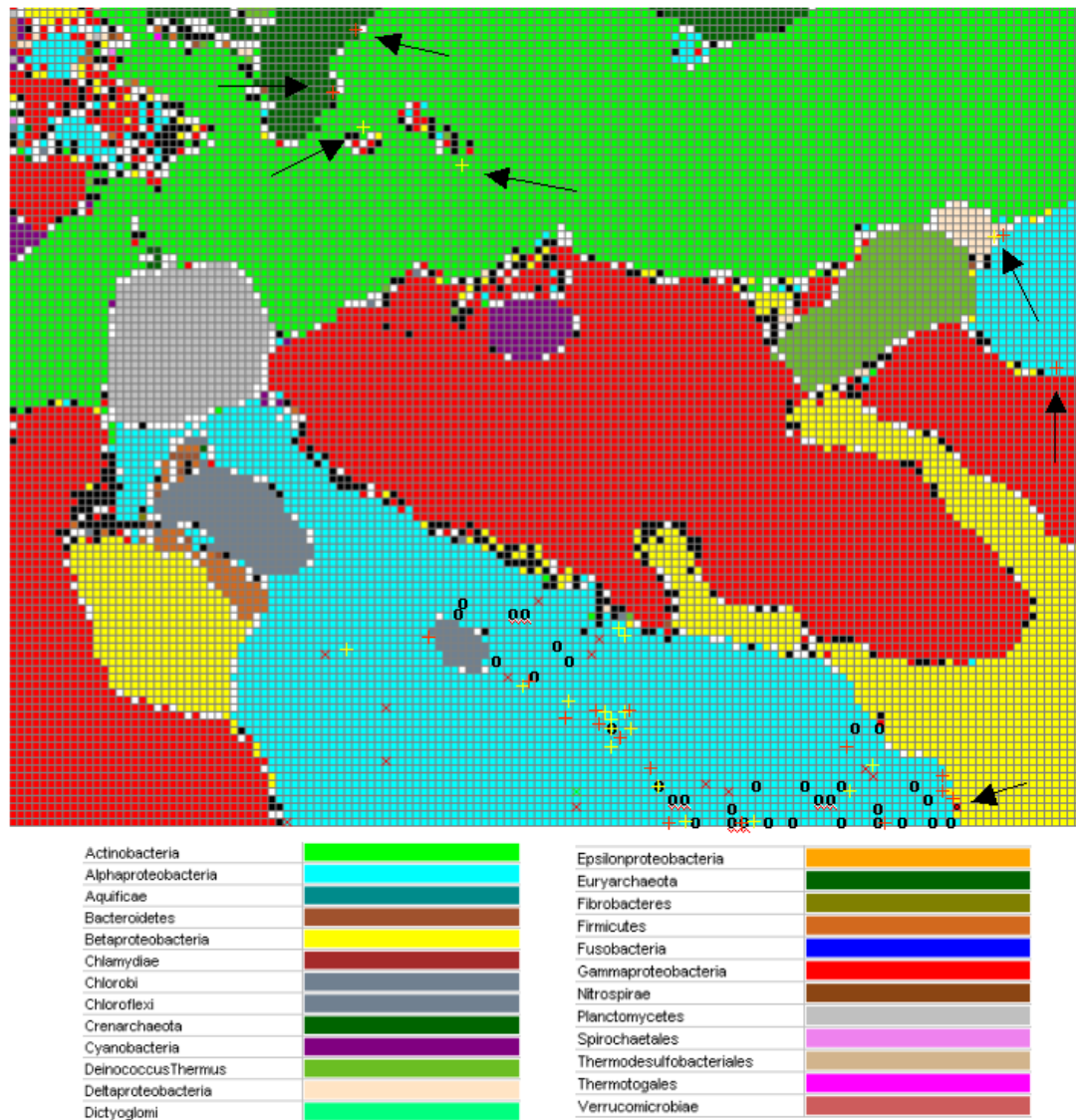


Figure 2. SOM for 5 kb fragments of public available prokaryotic genomes computed by the XanaGene Software. Symbols indicate the calculated positions of 5 kb genomic fragments of USC α (x), *M. acidiphila* (o), *Methylosinus* spp. (+,orange) and *Methylocystis* spp. (+,yellow).

Comparative analysis of genes encoding pMMO. General aspects. The *pmo* genes of both USC α and *M. acidiphila* are arranged in a single operon in the order *pmoCAB*. The comparison of predicted secondary structures, in particular of transmembrane helices (Fig. 4), with those reported previously for type I and type II MB (Gilbert et al., 2000; Ricke et al., 2004); (Stolyar et al., 1999) revealed a high degree of similarity (data not shown). For both, USC α and *M. acidiphila*, the same number of transmembrane helices were predicted for derived PmoCAB as follows (numbers are given in parenthesis): PmoA (6), PmoB (3), and PmoC (6).

M. acidiphila is a member of the *Alphaproteobacteria*. However, it branches in 16S rDNA-based phylogenies clearly distinct from the *Methylosinus/Methylocystis* group (Dedysh et al., 2001). The phylogenetic distinction is also reflected by both the *pmoCAB* identity values (data not shown) and the identity (similarity) values calculated at the amino acid level. For example, PmoCAB of *M. acidiphila* exhibited with that of *Methylocystis* sp. M (Gilbert et al., 2000) the following identity (similarity) values: PmoC: 64% (77%); PmoA: 69% (87%); and PmoB 57% (70%). The values computed for the gammaproteobacterial *M. capsulatus* (Bath) were constantly even 3 to 7 percent points lower. For comparison: similarity (identity) values calculated for *pmoCAB* of two type II MB are in the range of 89%-95% (81%-87%) (Gilbert et al., 2000); (Ricke et al., 2004).

In comparative analyses of *pmoA* sequences (507-bp fragment), USC α exhibited highest similarities to *M. acidiphila*. The same is true independent of whether the concatenated PmoCAB sequence or each individual polypeptide sequence was used for comparative analysis (Table 1). In accordance with the fact that PmoA carries the active center of pMMO, the highest degree of sequence conservation is found for *pmoA*/PmoA. The similarity (identity) of PmoCAB of USC α and *M. acidiphila* averages 81,5 % (69,4%) and thus is clearly below the similarity values within the type II MB (see above).

Phylogenetic analysis of complete PmoCAB /AmoCAB sequences. For the first time, the data set of fully sequenced *pmo* and *amo* operons enabled a phylogenetic analysis of the concatenated *pmoCAB* (PmoCAB) and *amoCAB* (AmoCAB) sequences. Regardless of the phylogenetic method (distance-based, maximum likelihood based), parameters and models (JK, JTT) used for treeing analyses, the *pmoCAB* (PmoCAB) sequences of USC α and *M. acidiphila* were grouped together. The computed tree topologies resembled those computed for *pmoA* with, however, one major exception: in *pmoA*(PmoA)-based phylogenies USC α

and *M. acidiphila* formed a common branch together with the conventional *pmoA1* (PmoA1) sequences of the *Methylosinus/Methylocystis* group. The novel *pmoA2* (PmoA2) sequences formed a separate cluster. The *pmoCAB2* gene cluster has recently been described to constitute an paralogous copy of the conventional *pmoCAB1* operon. In contrast to the PmoA-based phylogenies, the concatenated PmoCAB sequences of USC α and *M. acidiphila* branch clearly separated from a lineage defined, on the one hand, by PmoCAB1 of type II MB and, on the other hand, by PmoCAB2 of *Methylocystis* strain SC2 (Fig 4). In addition, trees were computed on the basis of only the PmoC, PmoA or PmoB sequences. While the branching patterns of the PmoA and PmoC trees resembled strongly those computed for the large data set of partial PmoA sequences, the PmoB phylogeny is nearly identical to those of the concatenated *pmoCAB* or PmoCAB sequences.

Table 1. Similarity (Identity) values of derived amino acid sequences of various pMMO and AMO

pmoCAB					pmoC				
	M. acid.	M.tricho.	M. cap.	N'spira	M. acid.	M.tricho.	M. cap.	N'spira	
USC α	81,5(69,4)	73,5(59,2)	66,8(51,0)	55,5(39,4)	80,8(67,9)	73,9(60,2)	64,7(53,8)	53,8(40,1)	USC α
M. acid.	-	74,7(61,5)	69,1(54,3)	58,2(40,9)	-	75,1(63,5)	64,2(55,0)	57,0(40,9)	M. acid.
M.tricho.		-	69,0(52,2)	56,6(39,2)		-	73,3(59,2)	56,3(39,0)	M.tricho.
M. cap.			-	62,4(43,1)			-	60,2(43,4)	M. cap.
N'spira				-				-	N'spira

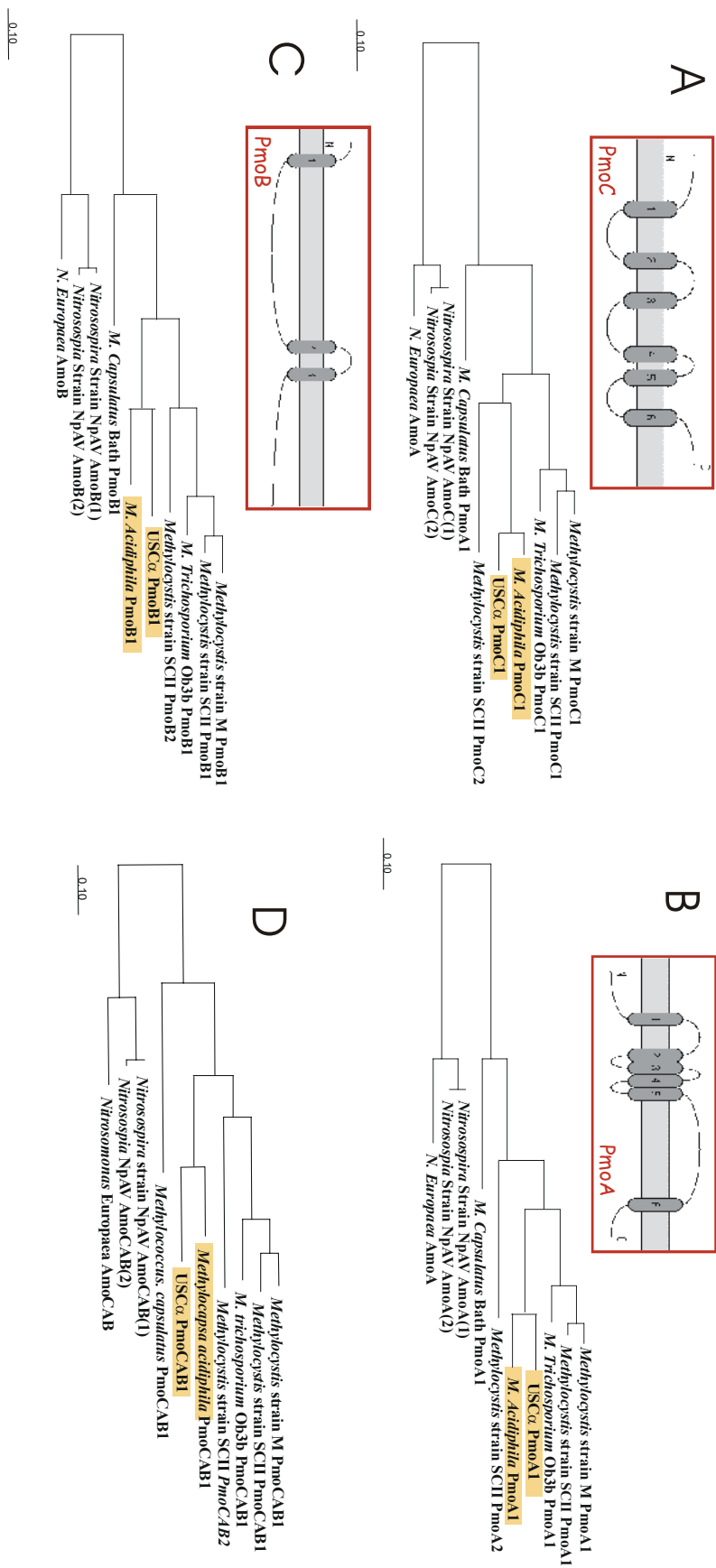
pmoA					pmoB				
	M. acid.	M.tricho.	M. cap.	N'spira	M. acid.	M.tricho.	M. cap.	N'spira	
USC α	88,1(77,5)	82,5(66,3)	75,2(56,4)	58,6(41,7)	76,3(64,3)	66,5(53,2)	62,3(45,2)	52,9(36,6)	USC α
M. acid.	-	84,2(68,0)	74,3(60,5)	59,4(40,6)	-	68,9(56,4)	69,6(50,1)	55,0(38,5)	M. acid.
M.tricho.		-	76,2(57,9)	59,1(40,2)		-	61,8(44,0)	53,2(36,9)	M.tricho.
M. cap.			-	66,4(43,1)			-	59,7(41,5)	M. cap.
N'spira				-				-	N'spira

M. acid = *M. acidiphila*, M. tricho = *M. trichosporium* OB3b, M.cap = *M. capsulatus* Bath, N'spira = *Nirosoospira* sp. NpAV (amoCAB1)

Figure 4. (next Page)

Boxes: Predicted transmembrane helices of USC α PmoCAB calculated by the toppred program. N-termini are predicted to be located in the cytosol. The topology strongly resembles those reported previously for PmoCAB of type X and type II MB. Helix 1 of PmoB is assumed to constitute a leader-sequence and therefore supposed to be absent in the active enzyme.

Phylogenetic trees computed based on the concatenated PmoCAB/AmoCAB sequences (d) or the corresponding individual polypeptides PmoC/AmoC (a), PmoA/AmoA (b) and PmoB/AmoB (c). Displayed trees were calculated by maximum likelihood algorithms but represent various used algorithms and models.



Identification of a conserved Orf downstream of *pmoB*. Directly downstream of *pmoB* on the USC α fragment an ORF (ORF45) was identified, which exhibited high similarity to genes found in the downstream region of *amoB* in autotrophic nitrifiers. Homologous genes had been reported for *N. europaea*, *Nitrosospira* sp. NpAV and *Nitrosococcus oceanii* ("*orf4*"), but also for *M. trichosporium* Ob3b ("*orfD*"). Besides *M. trichosporium* OB3b, no other type II MB had been shown to possess *orf4/orfD*. Also, no gene with significant similarities to ORF45 could be identified in the 101-kb fragment of *M. acidiphila*. However, Blast analysis against the complete genome of *M. capsulatus* (Bath) revealed a copy of this gene (MCA2130). This gene is located separately from the two *pmoCAB* operon copies of *M. capsulatus* (Bath). Based on both similarity values and phylogenetic analyses, the amino acid sequences derived from ORF45 (USC α) and ORF NE2060 (*N. europaea*) are most similar (57% similarity, 40% identity). Interestingly, ORF NE2060 is not located directly downstream of *amoB* but instead separated from *amoB* by the homologous gene NE2061 ("*orf4*"). Both upstream and downstream of ORF45, factor-independent terminators could be identified. In addition, a Shine-Dalgarno sequence was detected upstream of ORF45. However, no putative promoter region could be identified.

Discussion.

Retrieval of genomic fragments of USC α . To our knowledge, every attempt made so far to obtain pure cultures or enrichments of members of the USC α failed. We therefore applied metagenomic (or environmental genomic) techniques to obtain large genomic fragments of this putatively high affine MOB. The *pmoA* is the only gene for which sequence information of USC α is available. Consequently, a *pmoA*-targeted PCR assay was used for screening and identification of USC α -specific clone inserts. Results of previous studies suggested that cloning of complete *pmoCAB* gene clusters in *E. coli* using multi-copy vectors might not be possible due to toxic effects (Gilbert et al., 2000; Nguyen et al., 1998; Stolyar et al., 1999). In addition to obtaining large genomic fragments, using single copy Fosmid vectors allowed us to circumvent this problem.

Construction and screening of an environmental large insert genomic library resulted in two clones containing genomic fragments of representatives of USC α . Despite the fact that USC α was shown to preominate the methanotrophic population of various

acidic upland soils, previous studies had reported relative low abundances of USC α in regard to whole bacterial populations. Therefore this study is one of the first successful applications of the metagenomic strategy for the concerted retrieval of large genomic fragments of a low-abundant group of assigned uncultured members. Unfortunately, the two genomic fragments identified in the course of this survey belong to a single or two highly related genotypes, as they exhibited 100 % sequence similarity on the overlapping region. Because screening was performed with two different PCR assays, that both had been previously shown to be universal for *pmoA* of USC α (Holmes et al., 1996, Kolb et al., 2003), it is unlikely that identification of two identical *pmoA*-sequences was due to a screening artifact.

Based on phylogenetic trees constructed for partial *pmoA* sequences, *M. acidiphila* can be assumed to be the closest cultured relative of USC α . Consequently, the *pmoA* operon and flanking genomic regions of *M. acidiphila* were analysed for comparison. Assuming a genome size of 4 Mb, the content of the *M. acidiphila* BAC library corresponds to roughly 20 genome equivalents of *M. acidiphila*. Nevertheless, we were able to identify only 8 clones carrying the *pmoA* gene.

Analysis of the phylogenetic origin of USC α Apart from the *pmo* gene cluster, which is found only in MOB, 26 of 28 conserved ORFs exhibited highest similarities to members of the *Alphaproteobacteria*, mainly to *B. japonicum* and *R. palustris*. Moreover, gene clusters of at least two genes organised in the same order as in the analysed USC α genomic fragment were found identified in genomes of *Bradyrhizobiaceae*. No gene cluster was found to be conserved in any other genome if it was not identified also in the *B. japonicum* genome. Moreover, no other completely sequenced genome contained regions homologous to all conserved USC α gene clusters. Taken together, the high degree of structural conservation as well as the high similarity (identity) values clearly reflect a strong accordance of the analysed USC α fragment with genomes of members of the alphaproteobacterial *Bradyrhizobiaceae*, in particular the *B. japonicum* genome. It is unlikely that this is due to random gene shuffling. Because the homologous genes of *B. japonicum* as well as of *R. palustris* are distributed over the whole genomes, it is also unlikely that the genes were obtained as cluster by horizontal gene transfer, i.e. that the analysed fragment of USC α is not characteristic for the complete genome.

ORF16 is the only predicted gene for which no high similarity values to an alphaproteobacterial gene could be determined. It exhibited significant sequence

similarities only to a gene of *N. europaea*. ORF16 is localised on the only genomic fragment predicted by the SOM-based analysis to belong to the *Betaproteobacteria*. Together these results strongly indicate that ORF16 was acquired via horizontal gene transfer from a member of the *Betaproteobacteria*.

In addition to gene- or protein-based similarity studies, we applied a method focusing on the analysis of (taxon specific) genomic signatures, i.e. di- tri- and tetra-nucleotide distribution patterns. This technique has been widely discussed during the last years (Deschavanne et al., 1999; Edwards et al., 2002; Karlin et al., 1994; Karlin et al., 1992; Karlin et al., 1997; Knight et al., 2001; Oliver et al., 1993; Sandberg et al., 2001, Sandberg et al., 2003; Teeling et al., 2004). The signature analysis was performed with the program XanaMap, developed at the Xanagene Inc (<http://www.xanagen.com/>). XanaMap bases on the unsupervised neuronal network algorithm self-organizing map (SOM), which was intensively tested and demonstrated to recognize taxon- and even species-specific characteristics (key combinations of oligonucleotide frequencies). It therefore allows the taxon- or species-specific classification of genomic sequences without any regard on coding regions or sequence similarities (Abe et al., 2002; Abe et al., 2003). The taxonomic SOM-based affiliation was attained by comparison of 5 kb genomic subfragments of USC α against a map computed out of 5 kb genomic fragments of the public available complete sequenced genomes. The SOM-based analysis clearly confirmed the gene- and genome-based phylogenetic assignment of USC α to the *Alphaproteobacteria*. Interestingly, the analysed fragments of *M. acidiphila* genome as well as of type II MOB were placed in similar regions of the "alphaproteobacterial map area", indicating at least that this sequences are closely related within the *alphaproteobacteria*

The results of both comparative sequence analysis of single genes or gene clusters and SOM-based analysis provide final evidence that USC α has an alphaproteobacterial origin. The two methodologically independent approaches group USC α with members of the *Bradyrhizobiaceae*, in particular with *B. japonicum*. However, this is also true for *M. acidiphila*, despite the fact that, based on 16S rDNA data as well as NifHD phylogenies (Dedysh et al., 2002), *M. acidiphila* clearly bid more closely related to *Beijerinckia* spp. than to the *Bradyrhizobiaceae*. Therefore, the implications of the genera-related placement of USC α as well as *M. acidiphila* should not be overestimated. Due to the lack of available genomic information from other *Alphaproteobacteria* necessary for a more detailed phylogenetic placement of USC α (in particular of the genus *Beijerinckia*, as well as

sufficient data of type II MB) our results do not provide evidence that USC α is less related to *Beijerinckia* or type II MB than to the *Bradyrhizobiaceae*.

To gain a deeper insight into the relationship between USC α , *M. acidiphila* and representatives of the *Bradyrhizobiaceae*, we analysed the phylogenies of five genes identified on the genomic fragments of both USC α and *M. acidiphila*, as well as in the genomes of various other bacterial strains. The resulting phylogenies will not be discussed in detail, as the derived trees are not always in complete agreement with 16S rDNA-based phylogenies. This may be due to ancient events of horizontal gene transfer, but it is also possible that the degree of conservation of the respective amino acid sequences is not high enough to produce a meaningful phylogenetic relationship for highly divergent lineages. However, these phylogenies clearly reflect the common history of the organisms under study, including USC α , *M. acidiphila*, and members of the *Bradyrhizobiaceae* and therefore are in deep agreement with the genomic analyses described above. They also show consistently the closer relationship of the two analysed MB on one hand and on the other hand of the two *Bradyrhizobiaceae*. Again, due to the lack of sequence information neither type II MB nor members of the genus *Beijerinckia* could be included in the analysis. The inclusion of these organisms would have been of specific interest because the direct comparison of type II MB and of USC α still is only possible based on *pmoCAB* / *PmoCAB* data. Nonetheless, we have been able to show that the close relationship of *M. acidiphila* and USC α is supported not only by pMMO but also by the phylogenies of various other genes.

Comparative Analysis of *pmoCAB* gene clusters. MB belonging to the so far uncultured group USC α are of specific interest not only because they are highly abundant in various acidic upland soils but also because it has been assumed for a long time, that they might be able to oxidize methane at atmospheric concentrations. Consequently, a major aim of this study was a comparative analysis of the genes encoding pMMO. The presumed capability of USC α to oxidize methane at atmospheric concentrations might be reflected at the molecular/enzymatic level. It should therefore be investigated whether the derived polypeptide sequences possess any unusual properties or characteristics. The pMMO sequences of USC α and *M. acidiphila* were compared with those of five *pmo* operons of type II and type X MOB and with three homologous *amo* operons of autotrophic nitrifiers.

The *pmo* genes of both USC α and *M. acidiphila* are organised in a single gene cluster in the order *pmoCAB*, indicating that the structural organisation of the *pmo* operon is

highly conserved among all MOB analysed so far. However, we detected on the USC α genomic fragment directly downstream of *pmoB* an ORF (ORF45) for which no homologous gene could be identified on the genomic fragment of *M. acidiphila*. In contrast, a homologon was present in the corresponding genomic regions of *M. trichosporium* OB3b and the betaproteobacterial nitrifiers ("OrfD"). Interestingly, despite the fact that *pmoCAB* of USC α is based on both sequence similarity values and phylogenies clearly more closely related to other *pmo* operons than to *amo* operons, the homologous gene most similar to ORF45 was identified in *N. europaea* (NE2060). The similarity value between the amino acid sequences derived from ORF45 and NE2060 is in the same range than similarity values between USC α PmoCAB and *N. europaea* AmoCAB. Both pMMO and AMO are key enzymes of the respective metabolic pathways and therefore subject to strong purifying evolutionary selection. The conserved gene organisation clearly suggests that ORF45 and NE2060 are separated for the same evolutionary period than the adjacent monooxygenase operons are. It therefore can be concluded that the purifying selection pressures acting on either ORF45/ORFD or pMMO/AMO are similar strong.

As reported previously (Norton et al., 2002) *orf4* seems not to be part of the *amoCAB* operon as it is separated from *amoB* by a terminator region. Similarly, ORF45 seems not to be part of the *pmoCAB* operon in USC α because a factor-independent terminator is located between *pmoB* and ORF45. However, the high degree of conservation between ORF45 and NE2060 suggests functional importance. Indeed, a Shine-Dalgarno sequence could be identified upstream of ORF45 and a terminator is located downstream of it, whereas no putative promoter could be identified in the 100-bp region between *pmoB* and ORF45. If ORF45 is expressed, as indicated by the purifying selection pressure acting on it, there are only two possibilities for transcription: a) ORF45 is a facultative part of *pmoCAB* and regulated e.g. via antitermination or b) its promoter is localised in the gene *pmoB*.

The secondary structures predicted for the derived pMMO's of both USC α and *M. acidiphila* are highly similar. They agree well with the secondary structures reported previously for pMMO of other MOB, including gammaproteobacterial *M. capsulatus* Bath as well as alphaproteobacterial *M. trichosporium* OB3b and *Methylocystis* sp. strain M (Gilbert et al., 2000; Rieke et al., 2004; Stolyar et al., 1999). In summary, our findings provide further evidence that both the structural organization of *pmoCAB* and the secondary structure of derived proteins are highly conserved among all MOB, regardless of their phylogenetic positions.

Phylogenies of complete pMMO sequences. With the sequence information obtained in this study for *M. acidiphila* and USC α , and the recently analysed *pmoCAB1* and *pmoCAB2* of *Methylocystis* strain SC2, the data set of fully sequenced pMMO /AMO operons became large enough for a phylogenetic analysis of concatenated gene sequences. Compared to treeing analyses based only on partial *pmoA*, our approach sextuples the underlying phylogenetic information. Nevertheless, regardless the evolutionary models and algorithms used, the phylogenies always corresponded well to *pmoA*-based trees with the only exception that the USC α / *M. acidiphila* group branched separated while the recently reported pMMO2 of *Methylocystis* strain SC2 formed a common branch with PmoCAB1 of type II MOB. Because pMMO2 had been shown to exhibit unusual properties, this finding again demonstrates the unique position of USC α / *M. acidiphila* among the known MOB. The possibility that the separate branching of USC α / *M. acidiphila* is only due to the small data set could be excluded by testing a small PmoA data set consisting only of the PmoA sequences used for concatenated PmoCAB analysis. In addition, trees based only on either PmoB or PmoC were computed. It could be shown that the separate branching of USC α / *M. acidiphila* PmoCAB is mainly due to the inclusion of PmoB (Fig.4).

References

- Abe,T., Kanaya,S., Kinouchi,M., Ichiba,Y., Kozuki,T., and Ikemura,T. (2002) A novel bioinformatic strategy for unveiling hidden genome signatures of eukaryotes: self-organizing map of oligonucleotide frequency. *Genome Inform. Ser. Workshop Genome Inform.* **13**: 12-20.
- Abe,T., Kanaya,S., Kinouchi,M., Ichiba,Y., Kozuki,T., and Ikemura,T. (2003) Informatics for unveiling hidden genome signatures. *Genome Res.* **13**: 693-702.
- Chistoserdova,L., Vorholt,J.A., Thauer,R.K., and Lidstrom,M.E. (1998) C1 transfer enzymes and coenzymes linking methylotrophic bacteria and methanogenic Archaea. *Science* **281**: 99-102.
- Dedysh,S.N., Horz,H.P., Dunfield,P.F., and Liesack,W. (2001) A novel pmoA lineage represented by the acidophilic methanotrophic bacterium *Methylocapsa acidiphila* [correction of acidiphila] B2.[erratum appears in Arch Microbiol 2002 Feb;177(2):200]. *Archives of Microbiology* **177**: 117-21.
- Dedysh,S.N., Khmelenina,V.N., Suzina,N.E., Trotsenko,Y.A., Semrau,J.D., Liesack,W., and Tiedje,J.M. (2002) *Methylocapsa acidiphila* gen. nov., sp. nov., a novel methane-oxidizing and dinitrogen-fixing acidophilic bacterium from Sphagnum bog. *International Journal of Systematic & Evolutionary Microbiology* **52**: 251-61.
- Dedysh,S.N., Panikov,N.S., and Tiedje,J.M. (1998) Acidophilic methanotrophic communities from Sphagnum peat bogs. *Applied & Environmental Microbiology* **64**: 922-9.
- Dunfield,P.F., Liesack,W., Henckel,T., Knowles,R., and Conrad,R. (1999) High-affinity methane oxidation by a soil enrichment culture containing a type II methanotroph. *Applied & Environmental Microbiology* **65**: 1009-14.
- Gilbert,B., McDonald,I.R., Finch,R., Stafford,G.P., Nielsen,A.K., and Murrell,J.C. (2000) Molecular analysis of the pmo (particulate methane monooxygenase) operons from two type II methanotrophs. *Applied & Environmental Microbiology* **66**: 966-75.
- Hanson,R.S. and Hanson,T.E. (1996) Methanotrophic bacteria. *Microbiol. Rev.* **60**: 439-471.
- Holmes,A.J., Costello,A., Lidstrom,M.E., and Murrell,J.C. (1995) Evidence that particulate methane monooxygenase and ammonia monooxygenase may be evolutionarily related. *FEMS Microbiol. Lett.* **132**: 203-208.
- Kanaya,S., Kinouchi,M., Abe,T., Kudo,Y., Yamada,Y., Nishi,T. et al. (2001) Analysis of codon usage diversity of bacterial genes with a self-organizing map (SOM): characterization of horizontally transferred genes with emphasis on the *E. coli* O157 genome. *Gene* **276**: 89-99.
- Knief,C., Lipski,A., and Dunfield,P.F. (2003) Diversity and activity of methanotrophic bacteria in different upland soils. *Appl. Environ. Microbiol.* **69**: 6703-6714.

Kolb,S., Knief,C., Stubner,S., and Conrad,R. (2003) Quantitative detection of methanotrophs in soil by novel pmoA-targeted real-time PCR assays. *Appl. Environ. Microbiol.* **69**: 2423-2429.

Norton,J.M., Alzerreca,J.J., Suwa,Y., and Klotz,M.G. (2002) Diversity of ammonia monooxygenase operon in autotrophic ammonia-oxidizing bacteria. *Archives of Microbiology* **177**: 139-49.

Rabus,R., Gade,D., Helbig,R., Bauer,M., Glockner,F.O., Kube,M. et al. (2002) Analysis of N-acetylglucosamine metabolism in the marine bacterium *Pirellula* sp. strain 1 by a proteomic approach. *Proteomics*. **2**: 649-655.

Ricke,P., Erkel,C., Kube,M., Reinhardt,R., and Liesack,W. (2004) Comparative analysis of the conventional and novel pmo (particulate methane monooxygenase) operons from methylocystis strain SC2. *Appl. Environ. Microbiol.* **70**: 3055-3063.

Stolyar,S., Costello,A.M., Peeples,T.L., and Lidstrom,M.E. (1999) Role of multiple gene copies in particulate methane monooxygenase activity in the methane-oxidizing bacterium *Methylococcus capsulatus* Bath. *Microbiology* **145**: 1235-44.

3.3 Untersuchungen von *nifH* und *nifD* methanotropher Bakterien

Übersicht :

Gegenstand der nachfolgend reproduzierten Veröffentlichung waren die Erhebung und der Vergleich der für Schlüsselenzyme der N₂-Fixierung kodierenden Gene methanotropher Bakterien. Die Sequenzanalyse von 25 *nifH*- und 15 *nifD*-Sequenzen ermöglichte die Untersuchung der molekularen Grundlagen der N₂-Fixierung methanotropher Bakterien. Darüberhinaus konnten anhand dieser Sequenzen erstmals in kultivierungs-unabhängigen Studien erzielte *nifH*-Umwelt-Sequenzen den acidophilen methanotrophen Gattungen *Methylocapsa* und *Methylocella* zugeordnet werden.

Die Phylogenien sowohl der von *nifH* wie auch von *nifD* abgeleiteten Aminosäure-Sequenzen stehen in hoher Übereinstimmung mit der Phylogenie der 16S rRNA-Gene. Die einzigen Ausnahmen bilden die Verzweigung der Nif-Sequenzen von *M. capsulatus* innerhalb alphaproteobakterieller Sequenz-Cluster sowie die sehr hohe Ähnlichkeit der Sequenzen von *M. acidiphila* und heterotrophen *Alphaproteobacteria* der Gattung *Beijerinckia*. Die molekularen Daten legen die Aufnahme eines alphaproteobakteriellen *nif*-Genclusters durch *M. capsulatus* (*Gammaproteobacteria*) mittels horizontalen Gentransfers nahe. Dies korrespondiert mit der für gammaproteobakterielle Methanotrophe einzigartigen Fähigkeit von *M. capsulatus*, N₂ unter hohen O₂-Konzentrationen (O₂-Konzentrationen bis zu 15%) zu fixieren. Diese Eigenschaft ist ansonsten ein Charakteristikum der Typ II MB. Ebenso korreliert die hohe Sequenz-Ähnlichkeit der *nif*-Gene von *M. acidiphila* und *Beijerinckia* spp. mit der von keinem anderen MB gezeigten Fähigkeit zur N₂-Fixierung unter atmosphärischen O₂-Konzentrationen. Insgesamt konnte eine starke Korrespondenz von molekularen *nif*-Sequenz-Daten und physiologischen Daten der N₂-Fixierung gezeigt werden.

Anmerkung :

Die in der nachfolgend wiedergegebenen Veröffentlichung dargestellten physiologischen Daten wurden von Dr. S. N. Dedysh erhoben. Ebenfalls durch Dr. Dedysh erfolgte die Anzucht der meisten für diese Arbeit genutzten Stämme, sowie, in einigen Fällen, die Extraktion der genomischen DNA.

NifH and NifD phylogenies: an evolutionary basis for understanding nitrogen fixation capabilities of methanotrophic bacteria

Svetlana N. Dedysh,¹ Peter Ricke² and Werner Liesack²

Correspondence

Werner Liesack

liesack@staff.uni-marburg.de

¹Institute of Microbiology, Russian Academy of Sciences, Moscow 117312, Russia

²Max-Planck-Institut für terrestrische Mikrobiologie, Karl-von-Frisch-Str., D-35043 Marburg, Germany

The ability to utilize dinitrogen as a nitrogen source is an important phenotypic trait in most currently known methanotrophic bacteria (MB). This trait is especially important for acidophilic MB, which inhabit acidic oligotrophic environments, highly depleted in available nitrogen compounds. Phylogenetically, acidophilic MB are most closely related to heterotrophic dinitrogen-fixing bacteria of the genus *Beijerinckia*. To further explore the phylogenetic linkage between these metabolically different organisms, the sequences of *nifH* and *nifD* gene fragments from acidophilic MB of the genera *Methylocella* and *Methylocapsa*, and from representatives of *Beijerinckia*, were determined. For reference, *nifH* and *nifD* sequences were also obtained from some type II MB of the alphaproteobacterial *Methylosinus*/*Methylocystis* group and from gammaproteobacterial type I MB. The trees constructed for the inferred amino acid sequences of *nifH* and *nifD* were highly congruent. The phylogenetic relationships among MB in the NifH and NifD trees also agreed well with the corresponding 16S rRNA-based phylogeny, except for two distinctive features. First, different methods used for phylogenetic analysis grouped the NifH and NifD sequences of strains of the gammaproteobacterial MB *Methylococcus capsulatus* within a clade mainly characterized by *Alphaproteobacteria*, including acidophilic MB and type II MB of the *Methylosinus*/*Methylocystis* group. From this and other genomic data from *Methylococcus capsulatus* Bath, it is proposed that an ancient event of lateral gene transfer was responsible for this aberrant branching. Second, the identity values of NifH and NifD sequences between *Methylocapsa acidiphila* B2 and representatives of *Beijerinckia* were clearly higher (98.5 and 96.6 %, respectively) than would be expected from their 16S rRNA-based relationships. Possibly, these two bacteria originated from a common acidophilic dinitrogen-fixing ancestor, and were subject to similar evolutionary pressure with regard to nitrogen acquisition. This interpretation is corroborated by the observation that, in contrast to most other diazotrophs, *M. acidiphila* B2 and *Beijerinckia* spp. are capable of active growth on nitrogen-free media under fully aerobic conditions.

Received 20 June 2003

Revised 27 November 2003

Accepted 7 January 2004

INTRODUCTION

Currently recognized acidophilic methanotrophic bacteria (MB) are represented by two genera, *Methylocella* and *Methylocapsa*, which belong to the *Alphaproteobacteria* (Dedysh *et al.*, 2000, 2002). Similar to the traditionally known alphaproteobacterial MB, *Methylosinus* and *Methylocystis*, acidophilic MB utilize the serine pathway for carbon

assimilation. However, they possess several unique morphological and physiological characteristics, and, based on 16S rRNA phylogeny, *Methylocella* and *Methylocapsa* are evolutionarily more closely related to acidophilic heterotrophic bacteria of the genus *Beijerinckia* than to the *Methylosinus*/*Methylocystis* group.

Acidophilic MB inhabit diverse acidic environments, including wetlands and upland soils of the boreal zone and tundra (Dedysh *et al.*, 2001, 2003, 2004; Dunfield *et al.*, 2003). These environments, especially *Sphagnum* peat bogs, are characterized by extremely low concentrations of available nitrogen compounds (Richardson *et al.*, 1978; Mitsch & Gosselink, 1986). Thus, the ability to utilize N₂ as a nitrogen source is a necessary phenotypic trait for

Abbreviations: DOT, dissolved oxygen tension; MB, methanotrophic bacteria; NJ, neighbour-joining.

The GenBank accession numbers for the nearly complete 16S rRNA gene sequences and the partial *nifH* and *nifD* gene sequences reported in this paper are AJ563925–AJ563935, AJ563937–AJ563961, and AJ608687–AJ608701.

microbial inhabitants of these ecosystems, including acidophilic MB. Both *Methylocella* and *Methylocapsa* are capable of dinitrogen fixation. However, members of these two genera behave in a different way when grown in liquid nitrogen-free media. Under high partial pressures of oxygen, representatives of *Methylocella* usually show quite poor growth in nitrogen-free media. This is also typical for other dinitrogen-fixing MB such as *Methylococcus*, *Methylosinus* and *Methylocystis* (Murrell & Dalton, 1983; Dedys *et al.*, 2000; Auman *et al.*, 2001). Decreasing dissolved oxygen tension can enhance both growth and acetylene reduction activity of these MB. In contrast, *Methylocapsa acidiphila* is capable of exponential growth in liquid nitrogen-free media under both aerobic and micro-aerobic conditions (Dedys *et al.*, 2002). This growth capability of *Methylocapsa acidiphila* is very similar to that of *Beijerinckia*, which was one of the first bacteria described as being capable of fixing dinitrogen effectively (Alston, 1936; Starkey & De, 1939; Becking, 1999).

This study aimed to compare the molecular genetic basis for the nitrogen fixation capabilities of acidophilic MB with those of phylogenetically related heterotrophic nitrogen fixers. Our investigation was based on the comparative sequence analysis of the *nifH* and *nifD* genes, which encode

dinitrogenase reductase and the α -subunit of dinitrogenase: the key components of the nitrogenase enzyme complex. Public-domain databases currently contain about 2000 *nifH* and 300 *nifD* sequences. However, *nifH* sequences for MB have become available only recently, after the corresponding PCR protocols had been established (Auman *et al.*, 2001; Boulygina *et al.*, 2002), and no studies have been done on *nifD* genes in these bacteria. Thus, our study offers the first comparison of MB 16S rRNA phylogeny with phylogenies based on two different structural genes of nitrogenase. Moreover, the results obtained for growth experiments of representative MB and *Beijerinckia* spp. on nitrogen-free media agreed well with the principal conclusions drawn from the phylogenetic trees constructed for NifH and NifD.

METHODS

Bacterial strains and growth conditions. Bacterial strains used in this study are listed in Table 1. The set of acidophilic MB included four type strains (*Methylocapsa acidiphila* B2, *Methylocella palustris* K, *Methylocella silvestris* BL2 and *Methylocella tundrae* T4), three strains of *Methylocella palustris* isolated either from the *Sphagnum* peatlands of north-eastern Germany (strain H4) or from the tundra wetlands of Russia (strains Ch3 and Y5), and two strains of *Methylocella tundrae* (Y1, and Ch1) isolated from two different tundra wetland sites in northern Russia and described by Dedys

Table 1. Bacterial strains used in this study

Group	Strain	Accession no. in culture collection or source of isolation
Acidophilic MB	<i>Methylocapsa acidiphila</i> B2	DSM 13967 ^T , <i>Sphagnum</i> peat bog, western Siberia, Russia
	<i>Methylocella palustris</i> K	ATCC 700799 ^T , <i>Sphagnum</i> peat bog, western Siberia, Russia
	<i>Methylocella palustris</i> Ch3	Tundra wetland, Chukotka, northern Russia
	<i>Methylocella palustris</i> Y5	Tundra wetland, Yugorsk, northern Russia
	<i>Methylocella palustris</i> H4	Peat bog lake Kleine Fuchskuhle, northern Germany
	<i>Methylocella silvestris</i> BL2	DSM 15510 ^T , acidic forest soil, Marburg, Germany
	<i>Methylocella tundrae</i> T4	DSM 15673 ^T , tundra wetland, Vorkuta, northern Russia
	<i>Methylocella tundrae</i> Y1	Tundra wetland, Yugorsk, northern Russia
	<i>Methylocella tundrae</i> Ch1	Tundra wetland, Chukotka, northern Russia
	<i>Methylococcus capsulatus</i> Texas	NCIMB 11853 ^T
Type I MB	<i>Methylobacter luteus</i>	NCIMB 11914 ^T
	<i>Methylomicrobium album</i>	NCIMB 11123 ^T
	<i>Methylosinus trichosporium</i> OB3b	ATCC 35070 ^T
Type II MB	<i>Methylosinus trichosporium</i> SM6	River Saale, Wichmar, Germany*
	<i>Methylosinus sporium</i> SM27a	River Saale, Wichmar, Germany*
	<i>Methylosinus sporium</i> 44/2	Spring bog, Rügen, Germany*
	<i>Methylocystis echinoides</i>	IMET 10491 ^T
	<i>Methylocystis</i> sp. H9a	Lake Haussee, Feldberg, Germany*
	<i>Beijerinckia indica</i> subsp. <i>indica</i>	ATCC 9039 ^T
	<i>Beijerinckia indica</i> subsp. <i>lacticogenes</i>	DSM 1719 ^T
Other Alphaproteobacteria	<i>Beijerinckia mobilis</i>	DSM 2326 ^T
	<i>Beijerinckia derxii</i> subsp. <i>derxii</i>	DSM 2328 ^T
	<i>Beijerinckia derxii</i> subsp. <i>venezuelae</i>	DSM 2329 ^T
	<i>Rhodoblastus acidiphila</i>	ATCC 25092 ^T
	<i>Azorhizobium caulinodans</i>	DSM 5975 ^T
	<i>Bradyrhizobium japonicum</i>	DSM 30131 ^T

*Strains described by Heyer *et al.* (2002).

et al. (2004). The set of reference organisms included nine strains of type I and type II MB, and eight type strains of dinitrogen-fixing heterotrophic bacteria.

Methylocapsa acidiphila B2 was grown on half-strength (1:2) liquid nitrogen-free M1 mineral medium (Dedysh et al., 2002). The same medium, in agar form, supplemented with 250 mg KNO₃ l⁻¹, was used for cultivation of all strains of *Methylocella*. Other MB were cultivated on basic mineral salts (MS) medium (Whittenbury et al., 1970). This medium was either supplemented with 1 g KNO₃ l⁻¹, to give nitrate mineral salts (NMS) medium, or was used without addition of nitrate to give nitrate-free mineral salts (NFMS) medium. All MB cultures were cultivated under a gas headspace containing 20 % (v/v) methane. The liquid cultures were shaken at 120 r.p.m. and incubated at 24 °C, with the exception of *Methylococcus capsulatus* Texas, which was incubated at 37 °C. *Beijerinckia* strains were cultivated on half-strength liquid nitrogen-free M1 mineral medium supplemented with 0.2 % (w/v) glucose. *Azorhizobium caulinodans*, *Bradyrhizobium japonicum* and *Rhodoblastus acidiphila* were grown on media recommended by the DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH) and ATCC, respectively.

Growth experiments. Cultures of *Methylocapsa acidiphila* B2, type I MB *Methylobacter luteus*, type II MBs *Methylocystis echinoides* IMET 10491 and *Methylosinus trichosporium* OB3b, and the heterotrophic bacterium *Beijerinckia indica* subsp. *lacticogenes* were grown in 120 ml serum flasks containing 30 ml of the appropriate nitrogen-free medium. Half-strength M1 medium was used in the case of *Methylocapsa* and *Beijerinckia*, and NFMS medium in the case of *Methylobacter*, *Methylocystis* and *Methylosinus*. Growth was monitored by measuring OD₆₀₀. Serum flasks containing 20 ml NFMS medium, with a range of pO₂ values from 0.005 to 0.2 bar, were inoculated with the MB cells obtained by cultivation in nitrogen-free media. The required pO₂ in the headspace of the flasks was obtained by purging flasks with N₂ after inoculation and then injecting O₂ via a syringe. Methane (10 %, v/v) was added to all flasks, which were then incubated at 24 °C on a shaker at 120 r.p.m.. At intervals of 2–4 h, the flasks were examined for growth by measuring OD₆₀₀ and the concentration of oxygen in the headspace. The latter was determined using a Shimadzu 8A gas chromatograph fitted with a thermal conductivity detector and a Molecular Sieve 5A column. The specific growth rate calculated for each incubation period was correlated with the mean pO₂ value observed for a particular flask at the beginning and end of the incubation period.

PCR amplification and sequencing of the 16S rRNA, nifH and nifD genes. Genomic DNA was isolated from cultures grown in liquid media or on agar, using the SDS-based procedure described previously (Dedysh et al., 1998). PCR-mediated amplification of the 16S rRNA gene from positions 28 to 1491 (numbering according to the International Union of Biochemistry nomenclature for *Escherichia coli* 16S rRNA) was performed for five novel strains of *Methylocella*, five type strains of *Beijerinckia*, and *Methylococcus capsulatus* Texas, as described elsewhere (Dedysh et al., 2000). A partial stretch of the *nifH* gene was amplified from DNA samples using a combination of two previously developed degenerate oligonucleotide primers: forward primer F1 (5'-TAYGGNAARGGNGGNATYGGNAARTC-3'), designed by Boulygina et al. (2002), and reverse primer nifH-r (5'-ADNGC-CATCATYCTNCC-3'), designed by Zehr & McReynolds (1989). The thermal profile was as follows: initial denaturation for 30 s at 94 °C, followed by 35 cycles consisting of denaturation at 92 °C for 60 s, primer annealing at 55 °C for 60 s, and elongation at 72 °C for 60 s. The final elongation step was extended to 5 min.

A newly designed primer pair was used for amplification of *nifD* gene fragments. The forward primer nifD-f (5'-CAGGAAATCTA-CATCGTCATGTC-3') and the reverse primer nifD-r (5'-TCCCA-NGARTGCATCTGRCGGA-3') yielded an approximately 1900 bp

fragment that started at position 436 of the *Bradyrhizobium japonicum* USDA 110 *nifH* gene. The amplicons included the 3' region of the flanking *nifH* gene and most (1380 bp) of the *nifD* gene. The thermal profile used to amplify *nifD* gene fragments was as follows: initial denaturation for 30 s at 94 °C, followed by 35 cycles consisting of denaturation at 92 °C for 30 s, primer annealing at 63 °C for 30 s, and elongation at 72 °C for 90 s. The final elongation step was extended to 5 min.

Alternatively, the forward primer nifD-f (5'-GYGGYTGC GCCTA-YGCCGG-3') could be used in combination with nifD-r to amplify *nifD* gene fragments, yielding a 1130 bp fragment that started at position 227 of the *Bradyrhizobium japonicum* USDA 110 *nifD* gene (GenBank X01045). This assay represented a confirmatory test for the presence of *nifD* and covered the possibility that *nifH* and *nifD* did not cluster contiguously within the same operon.

All PCR reaction mixtures contained 0.5 µl template DNA, 5 µl 10 × reaction buffer, 1.5 mM MgCl₂, 200 µM of each dNTP, 0.3 mM of each primer, and 1.25 U *Taq* DNA polymerase (Promega). Amplification was carried out in a total volume of 50 µl in 0.2 ml reaction tubes. Based on a previous report on the occurrence of *nifH* genes among MB (Auman et al., 2001), *Methylosinus trichosporium* OB3b and *Methylomicrobium album* were chosen as positive and negative controls, respectively. The 16S rRNA, *nifH* and *nifD* gene amplicons were purified using QIAquick spin columns (Qiagen) and sequenced on an ABI Prism 377 DNA sequencer using BigDye terminator chemistry, as specified by the manufacturer (PE Applied Biosystems).

Data analysis. Based on sequence information deposited in public-domain databases or generated in the course of this study, we established databases for *nifH* and *nifD* sequences using the ARB program package (<http://www.arb-home.de>). The databases were checked manually for correct alignment. The alignments were used for phylogenetic analyses without making changes to possible errors in the public-domain *nifH* and *nifD* sequences or the inferred amino acid sequences. The newly obtained 16S rRNA gene sequences were added to a database of about 24 000 nearly complete 16S rRNA sequences. Phylogenetic analyses were carried out using PHYLIP (Felsenstein, 1989) and Tree Puzzle (Strimmer & von Haeseler, 1996). Multiple trees were constructed for both 16S rRNA genes and amino acid sequences inferred from *nifH* and *nifD*, using distance-based (neighbour-joining, NJ) and maximum-likelihood methods, including DNAML (16S rRNA), ProtML (NifH, NifD), and Tree Puzzle (16S rRNA, NifH, NifD), in combination with various models of evolution, such as JC (16S rRNA), PAM, JTT and WAG (NifH, NifD). The final tree topologies shown for NifH and NifD are consensus trees that take into consideration only those branch points which were confirmed by each of the treeing methods and models of evolution applied. In addition, the significance levels of interior branch points obtained in NJ analysis were determined by bootstrap analyses (1000 data resamplings). The overall identity values between pairs of 16S rRNA gene sequences and of inferred amino acid (NifH, NifD) sequences were calculated using the appropriate tools of ARB.

Codon usage tables for *Methylococcus capsulatus* Bath, as well as alphaproteobacterial reference organisms such as *Bradyrhizobium japonicum*, were obtained from the Kazusa codon usage database (<http://www.kazusa.or.jp/codon/>) (Nakamura et al., 2000). Alternatively, a table for *Methylococcus capsulatus* Bath was computed based on 57 coding sequences (>21 000 codons) obtained from GenBank. Deviations of *Methylococcus capsulatus* *nifH/nifD* codon usage from the mean species-specific values deposited in the codon usage tables were computed using the graphical codon usage analyser (<http://gcua.schoedl.de/>) and compared to those of genes encoding particulate methane monooxygenase (*pmoCAB*). Using the same on-line tool, the

codon usage patterns of various alphaproteobacterial organisms were compared with those computed for *Methylococcus capsulatus*.

RESULTS

Phylogeny of acidophilic MB based on the 16S rRNA gene

Comparative sequence analysis of the 16S rRNA gene showed that acidophilic MB of the genera *Methylocella* and *Methylocapsa*, along with acidophilic heterotrophic bacteria of the genus *Beijerinckia*, form a monophyletic cluster within the *Alphaproteobacteria* (Fig. 1). Trees constructed with different methods, including NJ and

maximum-likelihood (DNAML, Tree Puzzle), provided support for the coherence of this cluster (data not shown).

Eight strains of *Methylocella* used in this study represented three currently known species of this genus, *Methylocella palustris*, *Methylocella silvestris*, and *Methylocella tundrae*, and formed a defined subgroup within the *Methylocella-Methylocapsa-Beijerinckia* cluster. The strains of *Methylocella palustris* and *Methylocella tundrae* were isolated from acidic *Sphagnum*-dominated peatlands of the boreal zone and the tundra (Dedys *et al.*, 2000, 2004), while the third species, *Methylocella silvestris*, was obtained from acidic forest soil (Dunfield *et al.*, 2003). Despite the fact that the strains of each species of *Methylocella* were obtained from

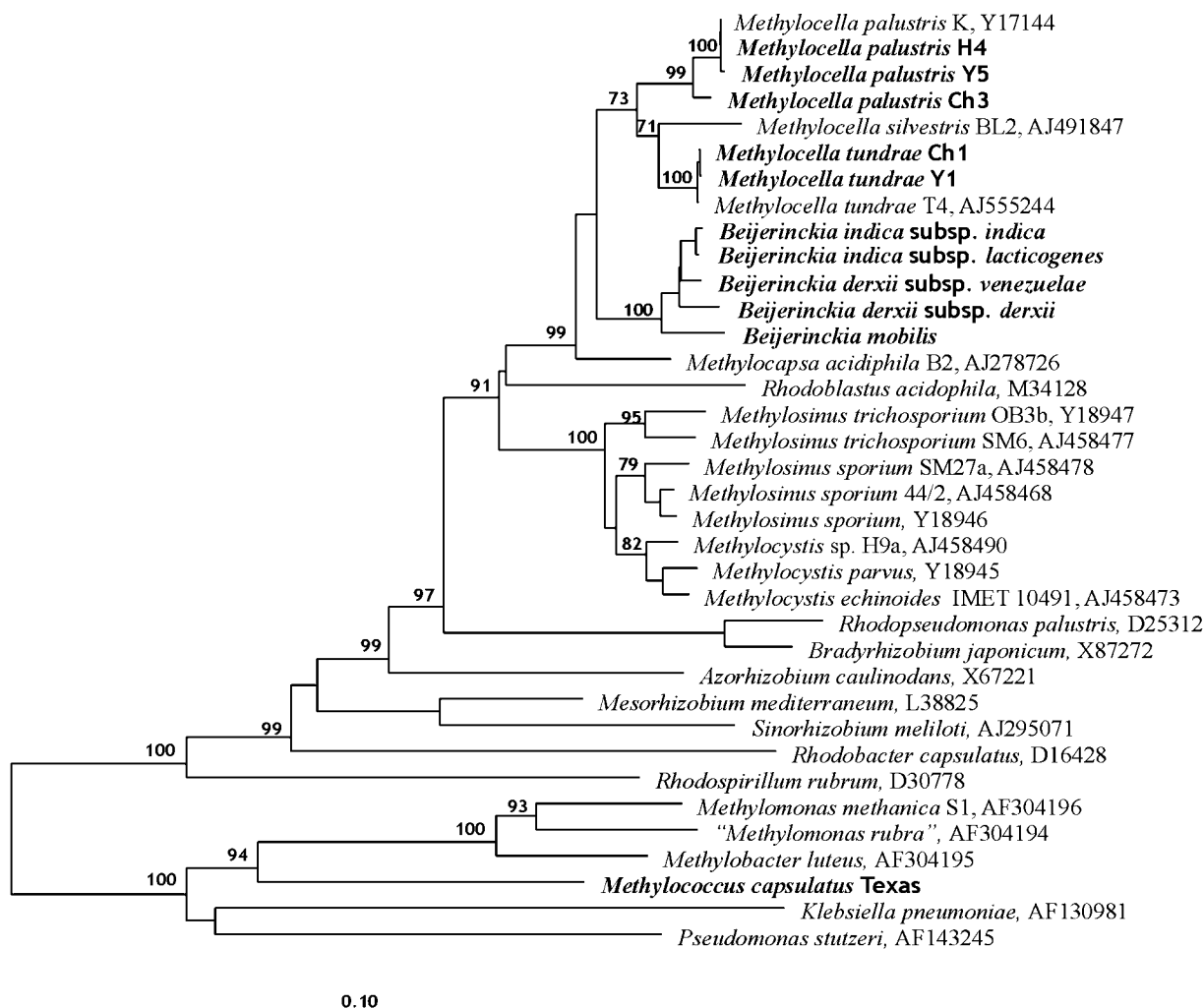


Fig. 1. 16S rRNA-based neighbour-joining tree showing the phylogenetic positions of acidophilic MB in relation to the dinitrogen-fixing heterotrophic bacteria of the genus *Beijerinckia* and to other representative dinitrogen-fixing members of the *Alphaproteobacteria*. The tree also includes 16S rRNA gene sequences of the gammaproteobacterial MB *Methylomonas methanica* S1, '*Methylomonas rubra*', *Methylobacter luteus* and *Methylococcus capsulatus* Texas, as well as some dinitrogen-fixing gammaproteobacterial heterotrophic bacteria. The numbers at the nodes indicate the percentage recovery in 1000 bootstrap resamplings. Micro-organisms for which sequence data were obtained in this study are shown in bold. GenBank accession numbers of the reference sequences used for tree construction are given. Scale bar, 0.1 substitution per nucleotide position.

geographically different locations (Table 1), the intraspecies 16S rRNA gene sequence identity was as high as 99.7–100%. For example, although strains K, Y5 and H4 of *Methylocella palustris* were isolated from western Siberia, the tundra zone of Russia, and northern Germany, respectively, they exhibited identical full-length 16S rRNA gene sequences. The genus *Methylocapsa* is currently represented by only one species, *Methylocapsa acidiphila*, and by the single strain B2. This might be due to the fact that these MB do not grow on agar media (Dedysh *et al.*, 2002), and are therefore very difficult to isolate.

Beijerinckia spp. are acidophilic, heterotrophic, dinitrogen-fixing bacteria. The four recognized species, *Beijerinckia indica*, *Beijerinckia mobilis*, *Beijerinckia derxii*, and *Beijerinckia fluminensis*, were taxonomically described half a century ago. However, the 16S rRNA gene sequence of *Beijerinckia indica* subsp. *indica* (Woese *et al.*, 1984) has until now been the only *Beijerinckia* sequence available in public-domain databases. We have determined the 16S rRNA gene sequences for four other representatives of this genus: *Beijerinckia indica* subsp. *lacticogenes*, *B. mobilis*, *B. derxii* subsp. *derxii* and *B. derxii* subsp. *venezuelae*. The 16S rRNA gene sequences of these species formed a phylogenetically distinct and tight cluster, and thus confirmed the assignment of these bacteria to the genus *Beijerinckia*. The sequence identities among the *Beijerinckia* spp. were in the range from 98.3 to 99.8% (Fig. 1), while the identities between members of the genus *Beijerinckia* and the acidophilic MB ranged from 96.0 to 97.3%.

Growth on nitrogen-free media

To compare the nitrogen-fixation ability of different MB in the absence of combined nitrogen, the growth on nitrogen-free media of the acidophilic MB *Methylocapsa acidiphila* B2, the type I MB *Methylobacter luteus*, and the type II MB *Methylocystis echinoides* IMET 10491 was monitored for 1 week (Fig. 2a). For comparison, the growth of *Beijerinckia indica* subsp. *lacticogenes*, on the same mineral medium used for *Methylocapsa acidiphila* B2, but with glucose added as the carbon source, was examined (Fig. 2b). The inoculum size was adjusted to achieve the same low initial OD₆₀₀ (0.03–0.04) for each culture analysed. The pO₂ in the headspace of the flasks was 0.18–0.19 bar. Under these conditions, both *Methylocapsa acidiphila* B2 and *Beijerinckia indica* subsp. *lacticogenes* showed exponential growth without a lag phase. The specific growth rates were 0.019 and 0.051 h⁻¹ for *Methylocapsa* and *Beijerinckia*, respectively. In contrast, the OD₆₀₀ for the culture of *Methylocystis echinoides* IMET 10491 remained nearly constant and a sharp decline of OD₆₀₀ was observed for *Methylobacter luteus*. After six days of incubation, the headspace pO₂ of the flasks inoculated with *Methylocystis* and *Methylobacter* was in the range 0.14–0.15 bar. To promote nitrogen fixation in these cultures, the flasks were purged with nitrogen and methane was reinjected into the headspace (shown by arrows in Fig. 2), giving a headspace pO₂ in the flasks of 0.01 bar. Both *Methylocystis echinoides*

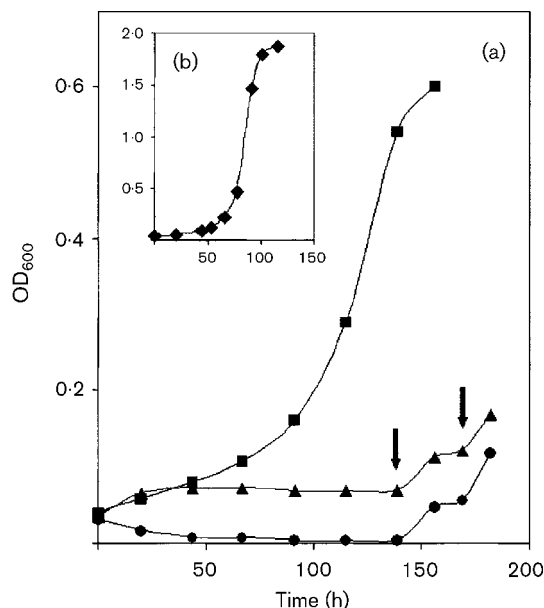


Fig. 2. Growth dynamics of three MB species (a) and of *Beijerinckia indica* subsp. *lacticogenes* (b) in batch cultures on nitrogen-free mineral media at high headspace pO₂ (0.18–0.19 bar). ■, *Methylocapsa acidiphila* B2; ▲, *Methylocystis echinoides* IMET 10491; ●, *Methylobacter luteus*. The pO₂ in the *Methylocystis echinoides* and *Methylobacter luteus* cultures was re-established during the experiment at 0.01 bar (arrows).

IMET 10491 and *Methylobacter luteus* responded to the reduction in pO₂ with immediate growth. However, only a relatively small increase of the OD₆₀₀ of both cultures was achieved, since O₂ was rapidly consumed by growing cells. Therefore, the injection of O₂ into the headspace of the flasks (pO₂ value of up to 0.01 bar) was repeated several times, and the cultures of the two MB species always showed the same growth response. Using this approach, we were able to obtain batch cultures of *Methylobacter luteus* and *Methylocystis echinoides* IMET 10491 of a relatively high OD₆₀₀, up to 0.35–0.45, on NFMS medium. In addition, we applied the same strategy to establish a batch culture of another type II MB, *Methylosinus trichosporium* OB3b, on NFMS medium, since this bacterium had already been used as a model organism to study the nitrogen-fixation capability of MB (Murrell & Dalton, 1983; Auman *et al.*, 2001). The cells from these three cultures were used to inoculate a number of flasks with a range of pO₂ values from 0.005 to 0.2 bar. This was done to assess the specific growth rate of different MB species on nitrogen-free medium as a function of headspace pO₂ (Fig. 3). The type I MB, *Methylobacter luteus*, was able to develop only at pO₂ values below 0.02 bar. In contrast, the two type II MB, *Methylosinus trichosporium* OB3b and *Methylocystis echinoides* IMET 10491, grew well under a wide range of headspace pO₂ values, up to 0.15–0.17 bar. However, growth of the type II MB at relatively high oxygen concentrations was possible only after multiple culture transfers

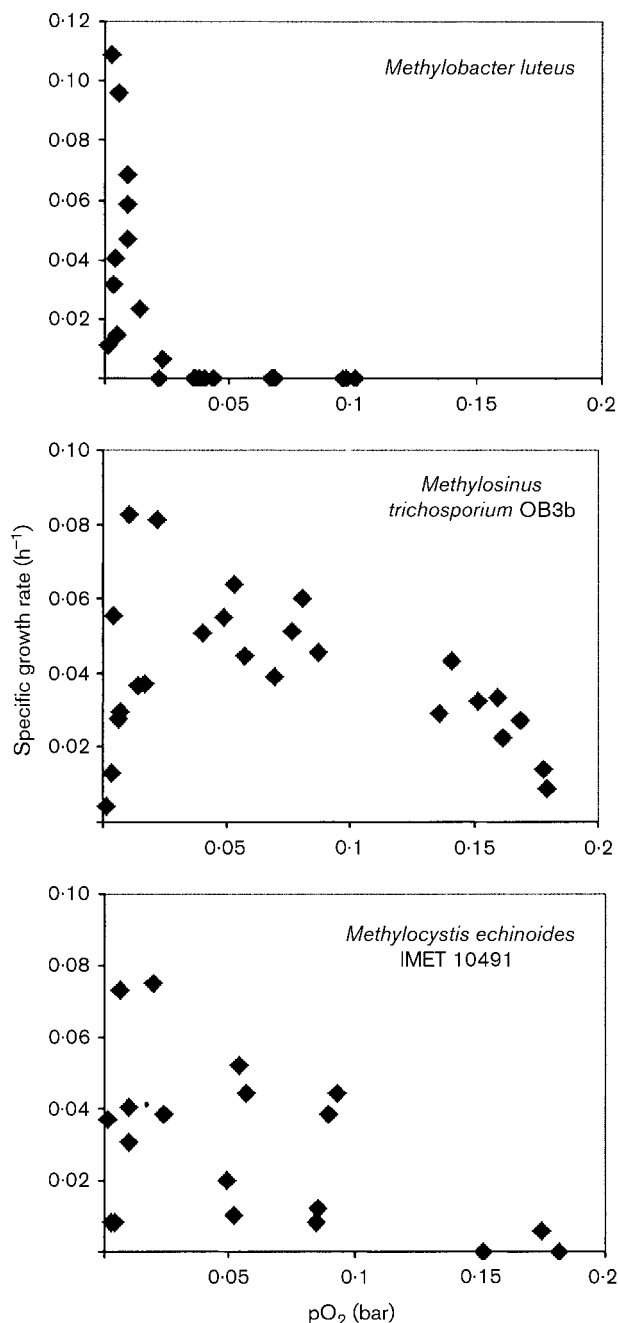


Fig. 3. Plot of the specific growth rates of different MB species versus pO₂ values in the headspace of the flasks. Cultures were grown in NFMS medium.

in NFMS medium. The highest specific growth rates on nitrogen-free medium, for both type I and type II MB, were recorded at low pO₂ values, below 0.02 bar.

PCR amplification of *nifH* and *nifD*

Two different protocols for *nifH* gene amplification from DNA of MB have been described. The protocol of Auman *et al.* (2001) was based on degenerate *nifH* primers (*nifH*-f

and *nifH*-r), designed by Zehr & McReynolds (1989). PCR with these primers yielded a *nifH* gene fragment of approximately 360 bp. Boulygina *et al.* (2002) designed another pair of *nifH*-targeted primers (F1 and R6), which enabled the retrieval of an approximately 450 bp *nifH* gene fragment. In this study, we used a combination of the forward primer F1 and the reverse primer *nifH*-r, which yielded a 453 bp PCR product. Amplification of *nifH* using this primer set showed consistent results for all MB and other related bacteria tested.

This work represents the first attempt to amplify *nifD* from the DNA of MB. Two pairs of newly designed primers enabled the retrieval of either an approximately 1130 bp fragment or an approximately 1900 bp fragment. The latter fragment, which consisted of the 3' region of *nifH* plus the almost complete *nifD* gene, was obtained from all the strains tested (Fig. 4b). Successful retrieval of these amplicons was possible only if *nifH* and *nifD* clustered contiguously within the same operon, suggesting that the structural organization of the *nif* genes in these strains is similar to that reported in most dinitrogen-fixing *Alphaproteobacteria* and *Gammaproteobacteria* (*nifHDK*) (Zehr *et al.*, 2003). For comparison, previously reported assays for *nifD* gene amplification yielded PCR products of only 450 to 830 bp (Ueda *et al.*, 1995b; Minerdi *et al.*, 2001; Parker *et al.*, 2002).

Until now, the database of *nifH* sequences from MB was limited to 26 sequences, of which only five sequences were derived from the type strains of recognized species. This study contributes another 25 *nifH* and 15 *nifD* sequences. This includes 17 *nifH* and 11 *nifD* sequences from MB, of which 16 were obtained from the type strains of MB species.

NifH- and NifD-based phylogenies of MB and related heterotrophic bacteria

General aspects. The topologies of the NifH- and NifD-based trees constructed in this study for MB and some other representatives of the *Alphaproteobacteria* and *Gammaproteobacteria* were very similar (Fig. 4). It should be noted that only 137 deduced amino acid residues were used for construction of the NifH tree, while the NifD tree is based on 356 deduced amino acid residues. Thus, the high congruity between the two trees suggests that the relatively short NifH sequence stretch provides sufficient phylogenetic information to reconstruct meaningful Nif-based trees for diazotrophic bacteria. In general, the clusters formed by the NifH and NifD sequences corresponded well with rRNA phylogeny. However, a few of the NifH and NifD sequences previously obtained from non-MB showed examples of an aberrant grouping. These were representatives of the *Betaproteobacteria*, namely *Herbaspirillum seropedicae*, and two strains of *Burkholderia*, free-living *Burkholderia fungorum* and symbiotic '*Candidatus Glomeribacter gigas*'. The NifH and NifD sequences of both species were affiliated with alphaproteobacterial-like Nif sequence types (compare Fig. 4a with 4b).

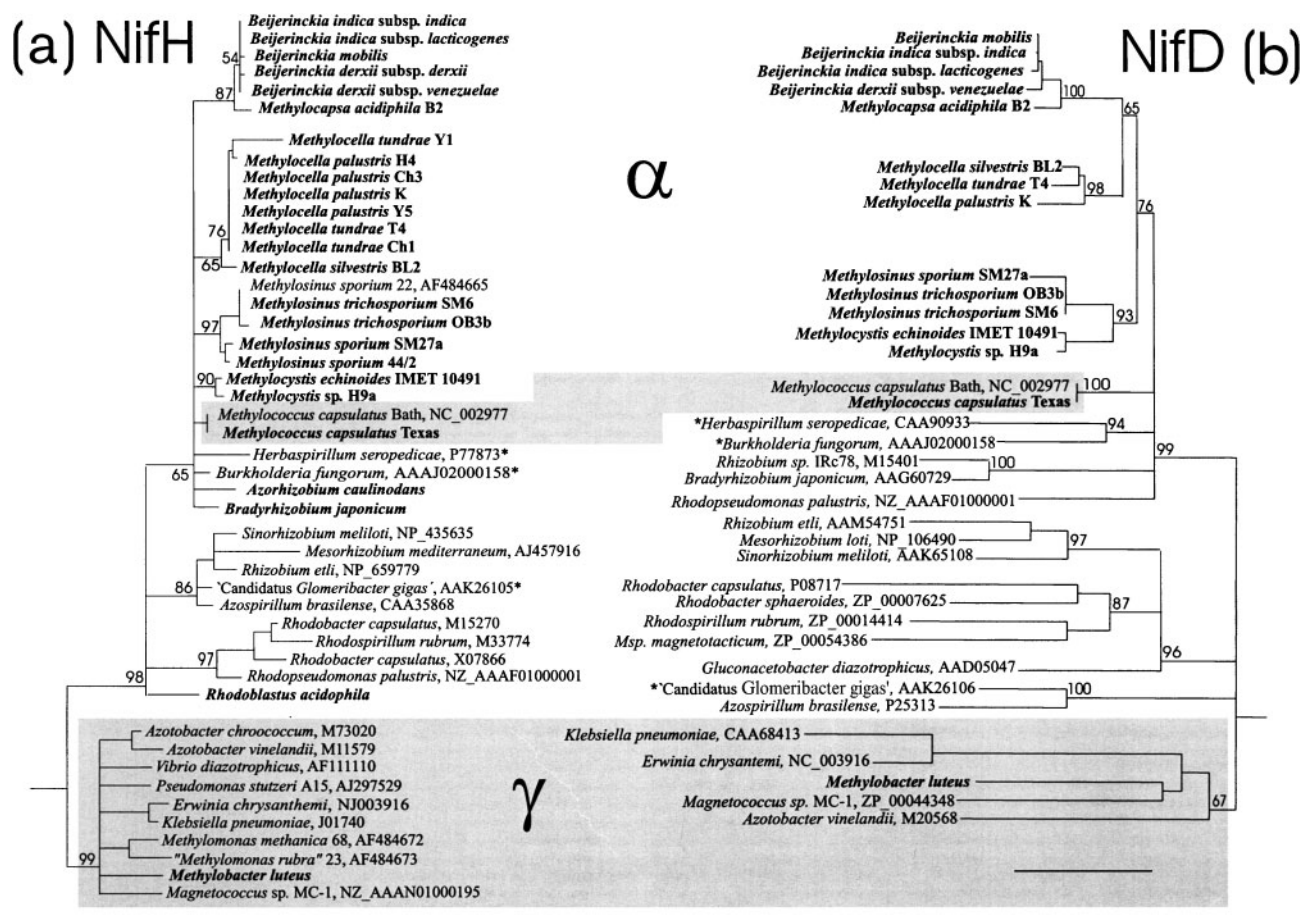


Fig. 4. Consensus trees of partial NifH sequences constructed based on 137 deduced amino acid sites (a) and partial NifD sequences constructed based on 356 deduced amino acid sites (b). The two trees show acidophilic MB in relation to *Beijerinckia*, the *Methylosinus*/*Methylocystis* group, gammaproteobacterial MB, and other representative dinitrogen-fixing members of the Alphaproteobacteria and Gammaproteobacteria. The assignment of bacterial taxa to either Alphaproteobacteria (white background) or Gammaproteobacteria (grey-shadowed) follows the 16S rRNA-based phylogeny, except for those marked *, which are Betaproteobacteria. Only branch points supported by both neighbour-joining and maximum-likelihood (ProtML, Tree Puzzle) methods are shown. The numbers at the nodes indicate the percentage recovery in 1000 bootstrap resamplings. Micro-organisms for which sequence data were obtained in this study are indicated in bold. The GenBank accession numbers of the NifH and NifD reference sequences used for construction of the respective trees are given. Scale bars, 0.1 substitution per amino acid position.

Alphaproteobacterial MB. In accordance with the 16S rRNA gene-based phylogeny, the NifH and NifD sequences from acidophilic MB and from the *Methylosinus*/*Methylocystis* group were affiliated with NifH and NifD sequences from other dinitrogen-fixing Alphaproteobacteria. In both NifH and NifD trees, these organisms formed distinct clusters composed of sequences from (i) *Methylocella*, (ii) *Beijerinckia*/*Methylocapsa acidiphila*, and (iii) *Methylosinus*/*Methylocystis*.

The NifH sequences from all eight representatives of the genus *Methylocella* grouped together (Fig. 4a), and the identity values between them varied between 98.5 and 100%. The NifH fragments obtained from strains of the

two peat-inhabiting species, *Methylocella palustris* and *Methylocella tundrae*, displayed complete sequence identity, while the NifH fragment from the upland soil-inhabiting species, *Methylocella silvestris*, was slightly divergent from the other sequence types of this cluster (1.5% sequence divergence). The *nifH* sequence-divergence among strains of *Methylocella* was slightly higher on the nucleic acid level (2.7 to 5.7%) than on the inferred amino acid level. The NifD sequences obtained from different species of *Methylocella* also formed a coherent cluster. The intracluster identity values ranged from 96.2 to 98.0%.

The NifH and NifD sequence identities within representatives of the genus *Beijerinckia* ranged from 97.7 to 100%,

and 98.5 to 99.1 %, respectively. Unexpectedly, the NifH fragment from *Methylocapsa acidiphila* B2 exhibited much higher identity values to the corresponding NifH fragments of *Beijerinckia* spp. (98.0–98.5 %) than to those of *Methylocella* spp. (90.8 %). In fact, the NifH fragment from *Methylocapsa acidiphila* B2 differed from the corresponding NifH fragments of the two subspecies of *Beijerinckia indica* by only two amino acid residues. A similar observation was made for NifD: the NifD fragment of strain B2 displayed sequence identities of 96.5–96.7 % to NifD fragments of *Beijerinckia* spp., and only 90.8–92.3 % sequence identities to NifD fragments of *Methylocella* spp.

The representatives of the *Methylosinus*/*Methylocystis* group were more divergent from each other than were the acidophilic bacteria. These organisms were separated in the NifH tree into two distinct lineages, characterized either by *Methylosinus* spp. or by *Methylocystis* spp., while their NifD sequences were clustered together into a common lineage. The identity values between NifH sequences of *Methylosinus* spp. and *Methylocystis* spp. ranged from 94.6 to 100 %, while the corresponding NifD values varied from 95.0 to 99.7 %. However, the strains selected for this study represent only a small proportion of the currently known strains of *Methylosinus* and *Methylocystis* (Heyer *et al.*, 2002), and thus our investigation might underestimate NifH and NifD divergence within this group.

Gammaproteobacterial MB. The NifH sequences from members of the MB genera *Methylomonas* and *Methylobacter* were assigned to lineages formed by NifH sequences of other *Gammaproteobacteria*. The identity values between the NifH sequences of alphaproteobacterial MB and those of *Methylomonas* and *Methylobacter* did not exceed 81.5 %. The identity values of NifD sequences from alphaproteobacterial MB to the corresponding fragment of *Methylobacter luteus* did not exceed 80.2 %, and thus provided further evidence for a clear separation of type I and type II MB in Nif-based phylogenies.

However, a striking disagreement between the tree topologies from, on the one hand, the 16S rRNA gene and, on the other hand, NifH and NifD, was identified for strains of the gammaproteobacterial MB *Methylococcus capsulatus*. The NifH and NifD sequences were grouped within clusters, characterized by alphaproteobacterial-like NifH and NifD sequence types, respectively. This disagreement was originally reported for NifH sequences from strains 114 and 115 of *Methylococcus capsulatus* (Boulygina *et al.*, 2002), while the taxonomic affiliation of these two strains was supported by partial 16S rRNA gene sequence data. To verify this aberrant grouping for other strains of *Methylococcus capsulatus*, we obtained sequence data for both the nearly complete 16S rRNA gene and partial *nifH* and *nifD* genes of *Methylococcus capsulatus* Texas (NCIMB 11853^T). In addition, we have exploited the data available for the *Methylococcus capsulatus* Bath genome (GenBank NC 002977). The NifH sequences of *Methylococcus capsulatus*

Texas and *Methylococcus capsulatus* Bath were identical to those of strains 114 and 115. The identity values of the four NifH sequences of *Methylococcus capsulatus* strains to those of the *Methylocella* and *Methylosinus*/*Methylocystis* groups were in the range 94.7 to 96.2 % and 93.2 to 95.4 %, respectively. The corresponding identity values calculated for NifD of *Methylococcus capsulatus* to those of the *Methylocella* and *Methylosinus*/*Methylocystis* groups ranged from 90.0 to 91.2 % and 90.5 to 91.4 %, respectively, while the identity to NifD of *Methylobacter luteus* was only 77.8 %.

Consideration of environmental NifH sequences

NifH sequences from acidophilic MB were used to perform a BLAST search against the GenBank database in order to identify closely related environmental clone sequences. A large number of NifH sequences retrieved from forest and agricultural soils of the boreal zone (Bürgmann *et al.*, 2004) displayed high identity values to the NifH sequences obtained in our study. The clone b1-WI1 (accession no. AY196397) was retrieved from acidic deciduous-forest soil of pH 4.6, and exhibited 99.1 % sequence identity to NifH from both *Beijerinckia indica* and *Methylocapsa acidiphila*. In this case, NifH-based phylogeny does not allow a reliable assignment of environmental *nifH* sequences into one of these two distinct acidophilic physiotypes. The sequence of the clone c1-HW3 (AY196440), which had been obtained from another acidic forest soil of pH 5.5, exhibited complete sequence identity to NifH from *Beijerinckia mobilis* and two subspecies of *Beijerinckia dextrii*. A group of environmental NifH sequences, including the clones b1-WI6, b1-WI2, f1-HW5, f1-PA10, f1-PA2, f1-PA4, f1-PA1, b1-GA6, b1-GA7, f1-HA1, and f1-HA2 (accession nos AY196383, AY196398, AY196454, AY196459, AY196460, AY196462, AY196458, AY196395, AY196396, AY196445, and AY196446, respectively) exhibited 99.1 % sequence identity to NifH sequences of *Methylocella palustris* and *Methylocella tundrae*. These clone sequences were retrieved from two acidic (pH 4.6 and 5.5) and one neutral (pH 6.8) forest soil, and from two agricultural soils (pH 7.2 and 7.5). The detection of these NifH sequences in soils of different acidity is not surprising, since *Methylocella* species are capable of growing in a pH range from 4.2 to 7.5. Two other environmental NifH sequences retrieved from deciduous forest soils of pH 6.8 and 5.5, clones b1-HA3-1 and f1-HW6 (AY196373 and AY196455), were closely related (98.2 % sequence identity) to NifH sequences from representatives of *Methylocella* and *Methylocystis*.

A BLAST search against the GenBank database was also performed using NifD sequences obtained in the study. This, however, did not yield any closely matching results, which might be due to the very limited number of environmental NifD clone sequences deposited in public-domain databases.

DISCUSSION

The ability to fix atmospheric dinitrogen is widespread in both *Archaea* and *Bacteria* (Postgate, 1987; Young, 1992; Martinez-Romero, 2000). The enzyme responsible for nitrogen fixation, nitrogenase, shows a high degree of conservation of structure and function across wide phylogenetic ranges (Dean & Jacobson, 1992). Nitrogenase is composed of two components, component I (dinitrogenase, or Fe-Mo protein), an $\alpha_2\beta_2$ tetramer encoded by the *nifD* and *nifK* genes, and component II (dinitrogenase reductase, or Fe protein), a homodimer encoded by the *nifH* gene (Fani *et al.*, 2000; Zehr *et al.*, 2003). The structural organization of the *nif* genes shows remarkable differences among nitrogen-fixing organisms (Zehr *et al.*, 2003). In most representatives of the *Alphaproteobacteria* and *Gammaproteobacteria*, *nif* genes are transcribed as a single transcriptional unit (*nifHDK* operon). However, in some *Alphaproteobacteria*, such as slow-growing rhizobia, the genes encoding nitrogenase are separated into two operons, so that the genes encoding dinitrogenase (*nifDK*) are transcribed from a promoter different from that of the structural gene for dinitrogenase reductase (*nifH*) (Yun & Szalay, 1984).

The cultivation-independent retrieval of *nif* sequences from different habitats has become a widely used approach to analyse the diversity of dinitrogen-fixing bacteria in ecosystems (Ueda *et al.*, 1995a, b; Zehr *et al.*, 1995, 1998; Widmer *et al.*, 1999; Zani *et al.*, 2000; Lovell *et al.*, 2001; Rösch *et al.*, 2002). However, with very few exceptions (Ueda *et al.*, 1995b), most of the environmental studies so far have focused on the retrieval and comparative sequence analysis of *nifH*. As a consequence, the currently existing public-domain databases are deficient with respect to *nifD*, while the rapidly expanding *nifH* dataset is dominated by sequences from micro-organisms which have not been cultivated. The lack of a sufficient number of *nif* gene sequences from cultured micro-organisms makes the interpretation of results obtained in environmental studies difficult and limits the clarity of the conclusions that can be drawn from phylogenetic treeing analyses. Our study contributes a representative set of *nifH* and *nifD* sequences from MB. These organisms are widely distributed in terrestrial and aquatic environments, and represent an important component of microbial assemblages in many nitrogen-depleted habitats. In addition, MB are known to differ in their nitrogen-fixation capabilities. This makes them an attractive target for comparative analysis of 16S rRNA- and *nif*-based phylogenies.

In general, *nifH* gene-based phylogenies have been shown to be in good agreement with the phylogenetic relationships derived from 16S rRNA gene sequences (Hennecke *et al.*, 1985; Young, 1992; Ueda *et al.*, 1995a; Zehr *et al.*, 1995). Comparison of phylogenies constructed for *nifH* and 16S rRNA genes from the same set of cultivated organisms showed little evidence of lateral gene transfer (Zehr *et al.*, 2003). However, a few cases of incongruent *nif*

and 16S rRNA phylogenies have also been reported, and some of these have been solidly supported by sequence data obtained for the whole *nif* operon. One example is the root-associated, dinitrogen-fixing species, *Herbaspirillum seropedicae*. Treeing analyses grouped the NifH and NifD sequences of this member of the *Betaproteobacteria* into a major clade, characterized mainly by *Alphaproteobacteria* (compare Fig. 4a with 4b) (Machado *et al.*, 1996). Other examples are the two nitrogen-fixing betaproteobacterial strains of *Burkholderia*, the free-living *Burkholderia fungorum* (whole-genome project NZ_AAAJ020 00001), and the symbiotic '*Candidatus Glomeribacter gigas*', which is a *Burkholderia* endosymbiont of the arbuscular mycorrhizal fungus *Gigaspora margarita* (Minerdi *et al.*, 2001). The NifH and NifD sequences of these organisms exhibit the same aberrant affiliation to alphaproteobacterial-like Nif sequence types as observed in *Herbaspirillum seropedicae*.

The genus *Burkholderia* represents an extremely heterogeneous group, including soil bacteria, plant-growth-promoting rhizobacteria, and human and plant pathogens. Thus, the number of *Burkholderia* spp. that exhibit an aberrant phylogeny of their *nifHD* gene sequences is still unknown, due to the limited number of strains for which data on both 16S rRNA and *nif* genes are available. Rösch *et al.* (2002) assessed the microbial diversity in acid forest soil by parallel characterization of environmental 16S rRNA and *nifH* gene sequences, using primer sets considered to be universal for *nifH*. It had been shown in previous studies that the *nifH*-targeted primers used by Rösch *et al.* (2002) amplify the corresponding gene fragment in a wide range of dinitrogen-fixing organisms, including members of the *Proteobacteria*, *Firmicutes* (representatives of both the *Actinobacteria* and the *Bacillus/Clostridium* group), and cyanobacteria. The 16S rRNA gene-based approach revealed an enormous bacterial diversity, while the *nifH* sequences were assigned mainly to a single cluster characterized by *Alphaproteobacteria*. One possible explanation for this finding is that many of the micro-organisms present did not possess *nifH*, or possessed *nifH* genes that were not targeted by the primer set used. Another hypothesis is that these micro-organisms have acquired alphaproteobacterial *nifH* gene-sequence types by lateral gene transfer. The high prevalence of alphaproteobacterial-like *nifH* gene-sequence types has also been reported in other environmental studies (Widmer *et al.*, 1999; Lovell *et al.*, 2001). In addition, most of the *nifH* sequences detected in the environment could not be assigned to any known dinitrogen-fixing bacterium. Thus the evolutionary significance of the lateral transfer of *nif* genes, as derived from the comparison of 16S rRNA gene- and Nif-based phylogenies, still remains unclear, because of the limited database of *nifH* and *nifD* sequences available for cultured bacteria.

However, the assignment of NifH and NifD sequences from strains of the gammaproteobacterial *Methylococcus capsulatus* to a major clade, characterized mainly by *Alphaproteobacteria*, is an additional example of disagreement

between the phylogenies constructed for the 16S rRNA gene and inferred *nif* sequences. This finding provides further support for the assumption that the lateral transfer of alphaproteobacterial *nif* gene-sequence types might have distributed effective nitrogen-fixation capabilities among diverse micro-organisms. Further analysis of the sequence data of the *Methylococcus capsulatus* Bath genome (GenBank NC 002977) confirmed that this MB has single copies of *nifH* and *nifD*. The NifH and NifD polypeptides are parts of the two components of nitrogenase, so that convergent evolution in both *nifH* and *nifD* of *Methylococcus* strains is unlikely and the phenomenon can more plausibly be explained by lateral gene transfer. However, the test for differences in codon usage between *nif* and *pmo* genes in *Methylococcus capsulatus* Bath failed to reveal any significant discrepancies. We also compared the codon-usage patterns calculated for *Methylococcus capsulatus* Bath with those of various diazotrophic *Alphaproteobacteria*, such as *Bradyrhizobium japonicum*. These patterns revealed a relatively high degree of similarity. A major determinant of codon usage is the genomic G + C content (Snyder & Champness, 2003). This content was very similar in *Methylococcus capsulatus* Bath and those *Alphaproteobacteria* used for comparison. As a consequence, even an evolutionarily recent incorporation of alphaproteobacterial-like *nif* genes into the *Methylococcus* genome by lateral gene-transfer might not be detectable by comparative analysis of codon-usage patterns. In addition, such an event of lateral *nifH* gene transfer could be ancient, so that the molecular traces of the event might have faded to the extent that they are undetectable.

Since the seminal studies of Whittenbury *et al.* (1970), it has been recognized that, although there are two basic groupings of MB, type I and type II, *Methylococcus* is an unusual genus because it shares properties of both types (Whittenbury, 1981; Whittenbury & Dalton, 1981). It has therefore been termed type X. The G + C content of DNA from representatives of *Methylococcus* (59–66 mol%) is closer to the corresponding values of type II MB (61–67 mol%) than to those of type I MB (49–60 mol%). However, all other phylogenetic assignments that have been made on the basis of comparative sequence analysis of the 16S rRNA gene, and genes encoding methane mono-oxygenase or methanol dehydrogenase, have consistently placed *Methylococcus* together with other representatives of gammaproteobacterial MB. Thus, the *nif* sequences provide the first molecular basis to highlight some of the physiological similarities observed between *Methylococcus* spp. and type II MB.

Interestingly, the ability to fix dinitrogen has been considered one of the major characteristics that distinguish representatives of the genus *Methylococcus* from other gammaproteobacterial MB. Murrell & Dalton (1983) used a combined approach, examining growth under a wide range of pO_2 values and using the acetylene reduction test to assess the dinitrogen-fixation capability of MB. This

study revealed that only alphaproteobacterial MB and *Methylococcus capsulatus* grow well in nitrogen-free liquid medium and actively reduce acetylene. Except for *Methylococcus* spp., gammaproteobacterial MB showed no growth in nitrogen-free medium, even if they were cultivated under microaerobic conditions. The findings of Murrell & Dalton (1983) correspond well with the results of our NifH- and NifD-based trees, in which all alphaproteobacterial MB plus *Methylococcus capsulatus* form a common cluster, clearly distinct from the gammaproteobacterial MB *Methylo-monas* and *Methylobacter*. Auman *et al.* (2001) demonstrated acetylene reduction activity in representatives of *Methylo-monas* and *Methylobacter*. However, growth of these organisms on nitrogen-free media has not been observed, nor has any strain ever been isolated using nitrogen-free enrichment conditions.

Thus, our investigation provides the first evidence that, besides *Methylococcus*, at least some representatives of type I MB are capable of growth in nitrogen-free media, though under a very limited range of conditions. It should be noted that in the experimental study of Murrell & Dalton (1983), a number of batch cultures of type I MB were also incubated under low pO_2 (0.01–0.02 bar) to achieve growth on nitrogen-free media. However, as pointed out above, the OD increase under these conditions is usually very small because the culture becomes limited by O_2 , and the OD soon declines back to zero. This transient growth cannot be detected if OD is measured at the beginning of the experiment and then at the end of a few days, as done by Murrell & Dalton (1983). A batch culture of type I MB of reasonably high OD was obtained in our experiments on nitrogen-free medium only by repeated injection of small amounts of oxygen to the headspace of the flasks. Alternatively, such a culture might be obtained using chemostat conditions, with O_2 as the growth-limiting substrate. This approach has been successfully applied to establish chemostat cultures of *Methylo-sinus trichosporium* OB3b and *Methylococcus capsulatus* Bath, and to show that nitrogen fixation in type II MB is less sensitive to pO_2 than in *Methylococcus* (Murrell & Dalton, 1983). The culture of *Methylo-sinus trichosporium* OB3b tolerated a dissolved oxygen tension (DOT) of 22 μM before wash-out occurred, while the maximum DOT tolerated by *Methylococcus capsulatus* Bath was 6 μM . However, routinely used aerobic batch-cultures contain about 250 μM of dissolved oxygen. Thus, it is not surprising that no type I or type II MB strain has ever been isolated using nitrogen-free enrichment conditions, and that growth on nitrogen-free media has never been demonstrated for these organisms. The sensitivity to O_2 of diazotrophic growth of MB might also explain why the maximum population density and activity of these bacteria in nitrogen-depleted environments always corresponds to sites with relatively low pO_2 values, such as the chemocline zone in water reservoirs and the layer below the water table in wetlands (Rudd *et al.*, 1976; Sundh *et al.*, 1994; Krumholz *et al.*, 1995). This pattern of MB distribution is usually explained by the availability of both

methane and oxygen. However, the autonomy in nitrogen acquisition, which can be realized only under a limited DOT range, is another explanation of the strict spatial localization of MB in natural environments.

The acidophilic heterotroph *Beijerinckia*, and the acidophilic MB *Methylocapsa*, represent the opposite case, where the genetic potential for dinitrogen fixation is realized under a wide range of pO₂ values. All species of *Beijerinckia* examined in this study were capable of growth on nitrogen-free media under fully aerobic conditions (results shown for *Beijerinckia indica* subsp. *lacticogenes* only). In contrast to some other well-studied free-living diazotrophic bacteria, such as *Azotobacter*, the mechanisms of dinitrogenase protection from oxygen in *Beijerinckia* have never been investigated. Protection of dinitrogenase from oxygen in *Azotobacter* occurs mainly through two mechanisms: a high respiratory activity that removes oxygen at the cell surface, and formation of alginate, which increases the viscosity of the culture broth and reduces the oxygen transfer rate from the gas phase to the aqueous phase (Sabra *et al.*, 2000). Like *Azotobacter*, strains of *Beijerinckia* also produce large amounts of polysaccharide material when grown in nitrogen-free media, so that the cultures become highly viscous. The cells of *Methylocapsa acidiphila* B2 are also covered by an extracellular polysaccharide matrix (Dedysh *et al.*, 2002), but this polysaccharide is not released into the medium. Thus, at present we cannot explain the mechanism of dinitrogenase protection in this acidophilic MB.

The almost complete identity of both NifH and NifD sequences in *Beijerinckia* and *Methylocapsa* demands some speculation about the evolutionary mechanisms that led to this phenomenon. One explanation would be that the two genera originated from a common acidophilic dinitrogen-fixing ancestor and subsequently experienced similar evolutionary selection pressures with regard to nitrogen acquisition. The *nifH* gene fragments of *Methylocapsa acidiphila* B2 and *Beijerinckia indica* subsp. *indica* exhibit a sequence divergence of 9.9%. The *nifD* gene fragments of these organisms display the same level of sequence divergence. However, most nucleotide sites that differ between the corresponding *nif* gene fragments in *Methylocapsa* and *Beijerinckia* are synonymous nucleotide substitutions and thus do not change the amino acid sequence. Apparently, selection pressure has maintained the primary structure of nitrogenase in these bacteria to enable their survival in nitrogen-depleted conditions.

The 16S rRNA gene-based phylogeny suggests, however, that members of the genus *Methylocella* also originated from the same common ancestor as *Methylocapsa* and *Beijerinckia*. Thus, although *Methylocella* and *Methylocapsa* have been isolated from the same acidic *Sphagnum* peat-bog environment, the pattern of *nif* gene evolution implies that *Methylocella* spp. possess a different niche-specialization for nitrogen from that of *Methylocapsa acidiphila* and *Beijerinckia* spp. (compare Figs 1 and 4). In contrast to *Methylocapsa acidiphila*, the representatives of

the genus *Methylocella* can utilize, in addition to inorganic nitrogen compounds, methylamines and some amino acids (Dunfield *et al.*, 2003; Dedysh *et al.*, 2004). Thus, the ability to fix dinitrogen is not a strictly obligatory trait for *Methylocella* spp., and the *nif* genes in these bacteria might have evolved in a different way.

ACKNOWLEDGEMENTS

This research was supported in part by the Russian Fund of Basic Research (grant no. 02-04-04003), the Program PCB of the Russian Academy of Sciences, the Deutsche Forschungsgemeinschaft [436 RUS 113/543/0-2(R)], and by the Competence Network Göttingen 'Genome research on bacteria' (GenoMik), financed by the German Federal Ministry of Education and Research (BMBF).

REFERENCES

- Alston, R. A. (1936). Studies on *Azotobacter* in Malayan soils. *J Agric Sci* **26**, 268–280.
- Auman, A. J., Speake, C. C. & Lidstrom, M. E. (2001). *nifH* sequences and nitrogen fixation in type I and type II methanotrophs. *Appl Environ Microbiol* **67**, 4009–4016.
- Becking, J. H. (1999). The genus *Beijerinckia*. In *The Prokaryotes*, 3rd edn. Edited by M. Dworkin and others. New York: Springer. <http://141.150.157.117:8080/prokPUB/index.htm>.
- Boulygina, E. S., Kuznetsov, B. B., Marusina, A. I., Tourova, T. P., Kravchenko, I. K., Bykova, S. A., Kolganova, T. V. & Galchenko, V. F. (2002). A study of nucleotide sequences of *nifH* genes of some methanotrophic bacteria. *Microbiology* (English translation of *Mikrobiologiya*) **71**, 500–508.
- Bürgmann, H., Widmer, F., von Sigler, W. & Zeyer, J. (2004). New molecular screening tools for the analysis of free-living diazotrophs in soil. *Appl Environ Microbiol* **70**, 240–247.
- Dean, D. R. & Jacobson, M. R. (1992). Biochemical genetics of nitrogenase. In *Biological Nitrogen Fixation*, pp. 763–784. Edited by G. Stacey, R. H. Burris & H. J. Evans. New York: Chapman & Hall.
- Dedysh, S. N., Panikov, N. S. & Tiedje, J. M. (1998). Acidophilic methanotrophic communities from *Sphagnum* peat bogs. *Appl Environ Microbiol* **64**, 922–929.
- Dedysh, S. N., Liesack, W., Khmelenina, V. N., Suzina, N. E., Trotsenko, Y. A., Semrau, J. D., Bares, A. M., Panikov, N. S. & Tiedje, J. M. (2000). *Methylocella palustris* gen. nov., sp. nov., a new methane-oxidizing acidophilic bacterium from peat bogs, representing a novel subtype of serine-pathway methanotrophs. *Int J Syst Evol Microbiol* **50**, 955–969.
- Dedysh, S. N., Derakhshani, M. & Liesack, W. (2001). Detection and enumeration of methanotrophs in acidic *Sphagnum* peat by 16S rRNA fluorescence in situ hybridisation, including the use of newly developed oligonucleotide probes for *Methylocella palustris*. *Appl Environ Microbiol* **67**, 4850–4857.
- Dedysh, S. N., Khmelenina, V. N., Suzina, N. E., Trotsenko, Y. A., Semrau, J. D., Liesack, W. & Tiedje, J. M. (2002). *Methylocapsa acidiphila* gen. nov., sp. nov., a novel methane-oxidizing and dinitrogen-fixing acidophilic bacterium from *Sphagnum* bog. *Int J Syst Evol Microbiol* **52**, 251–261.
- Dedysh, S. N., Dunfield, P. F., Derakhshani, M., Stubner, S., Heyer, J. & Liesack, W. (2003). Differential detection of type II methanotrophic bacteria in acidic peatlands using newly developed 16S rRNA-targeted fluorescent oligonucleotide probes. *FEMS Microbiol Ecol* **43**, 299–308.

- Dedysch, S. N., Berestovskaya, Y. Y., Vasylieva, L. V., Belova, S. E., Khmelenina, V. N., Suzina, N. E., Trotsenko, Y. A., Liesack, W. & Zavarzin, G. A. (2004). *Methylocella tundrae* sp. nov., a novel methanotrophic bacterium from acidic peatlands of tundra. *Int J Syst Evol Microbiol* **54**, 151–156.
- Dunfield, P. F., Khmelenina, V. N., Suzina, N. E., Trotsenko, Y. A. & Dedysch, S. N. (2003). *Methylocella silvestris* sp. nov., a novel methanotroph isolated from an acidic forest cambisol. *Int J Syst Evol Microbiol* **53**, 1231–1239.
- Fani, R., Gallo, R. & Lio, P. (2000). Molecular evolution of nitrogen fixation: the evolutionary history of the *nifD*, *nifK*, *nifE*, and *nifN* genes. *J Mol Evol* **51**, 1–11.
- Felsenstein, J. (1989). PHYLIP – phylogeny inference package (version 3.2). *Cladistics* **5**, 164–166.
- Hennecke, H., Kaluza, K., Thöny, B., Fuhrmann, M., Ludwig, W. & Stackebrandt, E. (1985). Concurrent evolution of nitrogenase genes and 16S rRNA in *Rhizobium* species and other nitrogen fixing bacteria. *Arch Microbiol* **142**, 342–348.
- Heyer, J., Galchenko, V. F. & Dunfield, P. F. (2002). Molecular phylogeny of type II methane-oxidizing bacteria isolated from various environments. *Microbiology* **148**, 2831–2846.
- Krumholz, L. R., Hollenback, J. L., Roskes, S. J. & Ringelberg, D. B. (1995). Methanogenesis and methanotrophy within a *Sphagnum* peatland. *FEMS Microbiol Ecol* **18**, 215–224.
- Lovell, C. R., Friez, M. J., Longshore, J. W. & Bagwell, C. E. (2001). Recovery and phylogenetic analysis of *nifH* sequences from diazotrophic bacteria associated with dead aboveground biomass of *Spartina alterniflora*. *Appl Environ Microbiol* **67**, 5308–5314.
- Machado, I. M., Yates, M. G., Machado, H. B., Souza, E. M. & Pedrosa, F. O. (1996). Cloning and sequencing of the nitrogenase structural genes *nifHDK* of *Herbaspirillum seropedicae*. *Braz J Med Biol Res* **29**, 1599–1602.
- Martinez-Romero, E. (2000). The dinitrogen-fixing bacteria. In *The Prokaryotes*, 3rd edn. Edited by M. Dworkin and others. New York: Springer. <http://141.150.157.117:8080/prokPUB/index.htm>.
- Minerdi, D., Fani, R., Gallo, R., Boarino, A. & Bonfante, P. (2001). Nitrogen fixation genes in an endosymbiotic *Burkholderia* strain. *Appl Environ Microbiol* **67**, 725–732.
- Mitsch, W. J. & Gosselink, J. G. (1986). *Wetlands*. New York: van Nostrand Reinhold.
- Murrell, J. C. & Dalton, H. (1983). Nitrogen fixation in obligate methanotrophs. *J Gen Microbiol* **129**, 3481–3486.
- Nakamura, Y., Gojobori, T. & Ikemura, T. (2000). Codon usage tabulated from the international DNA sequence databases: status for the year 2000. *Nucleic Acids Res* **28**, 292.
- Parker, M. A., Lafay, B., Burdon, J. & van Berkum, P. (2002). Conflicting phylogeographic patterns in rRNA and *nifD* indicate regionally restricted gene transfer in *Bradyrhizobium*. *Microbiology* **148**, 2557–2565.
- Postgate, J. (1987). *Fundamentals of nitrogen fixation (New Studies in Biology)*, 2nd edn. London: Edward Arnold.
- Richardson, C. J., Tilton, D. L., Kadlec, J. A., Chamie, J. P. M. & Wentz, W. A. (1978). Nutrient dynamics of northern wetland ecosystems. In *Freshwater Wetlands – Ecological Processes and Management Potential*, pp. 217–241. Edited by R. E. Good, D. F. Whigham & R. L. Simpson. New York: Academic Press.
- Rösch, C., Mergel, A. & Bothe, H. (2002). Biodiversity of denitrifying and dinitrogen-fixing bacteria in an acid forest soil. *Appl Environ Microbiol* **68**, 3818–3829.
- Rudd, J. W. M., Furutani, A., Flett, R. J. & Hamilton, R. D. (1976). Factors controlling methane oxidation in shield lakes: the role of nitrogen fixation and oxygen concentration. *Limnol Oceanogr* **21**, 357–364.
- Sabra, W., Zeng, A. P., Lunsdorf, H. & Deckwer, W. D. (2000). Effect of oxygen on formation and structure of *Azotobacter vinelandii* alginate and its role in protecting nitrogenase. *Appl Environ Microbiol* **66**, 4037–4044.
- Snyder, L. & Champness, W. (2003). *Molecular Genetics of Bacteria*, 2nd edn. Washington, DC: American Society for Microbiology.
- Starkey, R. L. & De, P. K. (1939). A new species of *Azotobacter*. *Soil Sci* **47**, 329–343.
- Strimmer, K. & von Haeseler, A. (1996). Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Mol Biol Evol* **13**, 964–969.
- Sundh, I., Nilsson, M., Granberg, G. & Svensson, B. H. (1994). Depth distribution of microbial production and oxidation of methane in northern boreal peatlands. *Microb Ecol* **27**, 253–265.
- Ueda, T., Suga, Y., Yahiro, N. & Matsuguchi, T. (1995a). Remarkable N₂-fixing bacterial diversity detected in rice roots by molecular evolutionary analysis of *nifH* gene sequences. *J Bacteriol* **177**, 1414–1417.
- Ueda, T., Suga, Y., Yahiro, N. & Matsuguchi, T. (1995b). Genetic diversity of N₂ fixing bacteria associated with rice roots by molecular evolutionary analysis of a *nifD* library. *Can J Microbiol* **41**, 235–240.
- Whittenbury, R., Phillips, K. C. & Wilkinson, T. F. (1970). Enrichment, isolation and some properties of methane-utilizing bacteria. *J Gen Microbiol* **61**, 205–218.
- Whittenbury, R. (1981). The interrelationship of autotrophy and methylophony as seen in *Methylococcus capsulatus* (Bath). In *Microbial growth on C1 compounds*, pp. 181–190. Edited by H. Dalton. London: Heyden.
- Whittenbury, R. & Dalton, H. (1981). The methylophytic bacteria. In *The Prokaryotes*, pp. 894–902. Edited by M. P. Starr and others. Heidelberg: Springer.
- Widmer, F., Shaffer, B. T., Porteous, L. A. & Seidler, R. J. (1999). Analysis of *nifH* gene pool complexity in soil and litter at a Douglas fir forest site in the Oregon Cascade Mountain range. *Appl Environ Microbiol* **65**, 374–380.
- Woese, C. R., Stackebrandt, E., Weisburg, W. G. & 8 other authors (1984). The phylogeny of purple bacteria: the alpha subdivision. *Syst Appl Microbiol* **5**, 315–326.
- Young, J. P. W. (1992). Phylogenetic classification of nitrogen-fixing organisms. In *Biological Nitrogen Fixation*, pp. 43–86. Edited by G. Stacey, R. H. Burris & H. J. Evans. New York: Chapman & Hall.
- Yun, A. C. & Szalay, A. A. (1984). Structural genes of dinitrogenase and dinitrogenase reductase are transcribed from two separate promoters in the broad host range cowpea *Rhizobium* strain Irc78. *Proc Natl Acad Sci U S A* **81**, 7358–7362.
- Zani, S., Mellon, M. T., Collier, J. L. & Zehr, J. P. (2000). Expression of *nifH* genes in natural microbial assemblages in Lake George, New York, detected by reverse transcriptase PCR. *Appl Environ Microbiol* **66**, 3119–3124.
- Zehr, J. P. & McReynolds, L. A. (1989). Use of degenerate oligonucleotides for amplification of *nifH* gene from the marine cyanobacterium *Trichodesmium thiebautii*. *Appl Environ Microbiol* **55**, 2522–2526.
- Zehr, J. P., Jenkins, B. D., Short, S. M. & Steward, G. F. (2003). Nitrogenase gene diversity and microbial community structure: a cross-system comparison. *Environ Microbiol* **5**, 539–554.

Zehr, J. P., Mellon, M., Braun, S., Litaker, W., Steppe, T. & Paerl, H. W. (1995). Diversity of heterotrophic nitrogen fixation genes in a marine cyanobacterial mat. *Appl Environ Microbiol* **61**, 2527–2532.

Zehr, J. P., Mellon, M. T. & Zani, S. (1998). New nitrogen-fixing microorganisms detected in oligotrophic oceans by amplification of nitrogenase (*nifH*) genes. *Appl Environ Microbiol* **64**, 3444–3450.

3.4 tRFcut : *in silico* tRFLP-Analyse in ARB

Übersicht :

Das Software-Paket ARB ist eine modulare, integrierte Umgebung für die phylogenetische Analyse molekularer Sequenzdaten. Ein zusätzliches Modul wurde programmiert, das die *in silico* tRFLP-Analyse von in ARB verwalteten Sequenz-Datensätzen ermöglicht. tRFcut ermöglicht die schnelle Zuordnung experimentell gewonnener tRFs zu bekannten Sequenzdaten und unterstützt das intelligente Design neuer tRFLP-Ansätze. Die wichtigsten Vorzüge von tRFcut gegenüber herkömmlichen tRFLP-Analyse-Programmen sind: i) nahezu unbeschränkte Kombinationsmöglichkeit beliebiger Primer und Restriktionsenzyme, ii) Anwendbarkeit nicht nur auf Sequenzen der 16S rRNA-Gene, sondern auf beliebige Datensätze, iii) Integration der Daten in eine effiziente Software-Umgebung und somit iv) Kombination der Berechnungen mit phylogenetischen Analysen sowie v) Möglichkeit der Analyse von Sequenzen mit fehlender Primerbindestelle.

Anmerkung :

Die in der nachfolgenden Veröffentlichung dargestellten experimentellen tRF-Daten wurden von Dr. G. Braker erhoben. Die analysierten DNA-Extrakte wurden von Dr. S. Kolb zur Verfügung gestellt. Beide waren als Tester in die Entwicklung von tRFcut involviert.

Application of a newly developed ARB-integrated *in silico* T-RFLP tool reveals the dominance of a novel *pmoA* cluster in a forest soil

Peter Ricke, Steffen Kolb, and Gesche Braker

Accepted for publication in Applied and Environmental Microbiology

Abstract

TRF-CUT, an ARB-implemented tool, was developed to predict *in silico* the terminal restriction fragments of aligned SSU rDNA or functional gene sequences. Application of this new tool to perform directed T-RFLP analysis on *pmoA* products obtained from a forest soil revealed that novel Cluster I methanotrophic bacteria were dominant.

T-RFLP (terminal restriction fragment length polymorphism) is a method commonly applied to study the structure of microbial communities (1, 2, 6, 14). PCR products of 16S rDNA or of functional genes end-labelled with a fluorescent dye are cleaved with a site-specific restriction endonuclease to obtain genetic fingerprints of microbial communities. Recently some online tools, e.g. TAP-TRFLP (12), torast (<http://www.torast.de>), and MiCA (<http://mica.ibest.uidaho.edu/>) have been developed to perform *in silico* hydrolysis of 16S rRNA sequences. However, analysis software predicting T-RFs from databases of functional marker genes (e.g. *pmoA*, *nirK*, *nifH*), which are used to describe the structure of functional microbial groups, is not available.

The freely available software ARB (11) is a graphics-oriented program including a database of SSU rDNA sequences and various sequence analysis tools. Since ARB is widely used for the phylogenetic analysis of strain-specific and environmental sequence data it was considered worthwhile to integrate an *in silico* T-RFLP analysis tool (TRF-CUT) into this software package. The upgrade with TRF-CUT is simple and does not alter the ARB source code. TRF-CUT and a manual file can be obtained from <http://www.uni-marburg.de/mpi/downloads/trfcut/trfcut.zip>.

In addition to the SSU RNA database TRF-CUT is applicable to any set of nucleic acid sequences aligned in ARB. Specification of parameters such as primer position and restriction enzyme allows the generation of individual T-RF commands that can be stored permanently. T-RFs are even predicted for sequences lacking a primer region, since the algorithm uses absolute alignment positions to determine the priming region instead of searching for individual primer sequences. This feature distinguishes TRF-CUT from any other T-RF prediction tool available to date, and is of special interest because priming regions are frequently omitted from sequences submitted to public databases, or may contain mismatches to the primer. In these cases these sequences would remain undetected by an *in silico* analysis, although they can be obtained by PCR. Furthermore, TRF-CUT indicates any species that have incomplete sequences between the primer region and the first restriction site, thus enabling the user to manually check or to exclude them from analysis if desired. T-RFs can be calculated with TRF-CUT for any restriction endonuclease with an unambiguous restriction site provided by the Rebase database (<http://rebase.neb.com/rebase/rebase.html>). Results of TRF-CUT are stored in ARB data fields and therefore can be used to perform searches and manipulations, and can

be combined for multiple enzyme digests. Results can be exported as lists e.g. for downstream analysis with the online tool PAT (8) or can be displayed at the tips of the phylogenetic trees, thus allowing the direct assignment and visualization of T-RFs specific to phylogenetic groups (Fig 1A). In conclusion, TRF-CUT is suitable for i) T-RF prediction based on various enzymes to select enzymes with high phylogenetic resolution and ii) assigning T-RFs from experimental T-RFLP data to potentially corresponding sequences in the database.

As an example we applied TRF-CUT to evaluate the community structure of methanotrophic bacteria (MB) in a pH-neutral upland soil located near Göttingen, Germany. Methanotrophic bacteria play a crucial role in the global carbon cycle. The initial oxidation of methane to methanol is catalysed by either a particulate (pMMO) or a soluble (sMMO) methane monooxygenase. The particulate, membrane-bound form, pMMO, occurs in almost all known MB and is homologous to the ammonium monooxygenase (AMO) of ammonia-oxidizing bacteria (AOB). The *pmoA* gene encodes the catalytic center of pMMO and is widely used as a functional marker in environmental studies (4, 9, 10). Its application to study MB communities in a variety of methane-consuming upland soils indicated that two novel *pmoA* sequence clusters USC α and USC γ were dominant (7, 9). Recently, novel MB belonging to the α Proteobacteria but harboring *pmoA* sequences only distantly related to known *pmoA* genes (here referred to as 'Cluster I') were isolated from a tundra soil (13). However, in contrast to USC α and USC γ , information on the distribution of this cluster is very limited.

Community analyses of MB by T-RFLP have been performed mainly by *MspI* hydrolysis of Fam-labeled *pmoA* PCR products (6), which were amplified using the primers A189f and A682r (5). In our study an *MspI* digest of *pmoA* amplified from the Göttingen forest soil sample resulted in four major peaks with one dominant T-RF of 80 bp (57%) (Fig. 1B). Application of TRF-CUT to a database including all public-domain *pmoA* and *amoA* sequences (June 2004) revealed that the 80-bp T-RF was indicative for an environmental sequence from an upland soil (E33b-a, Cluster I) and for *Methylobacter* sp. LW12. A 79-bp T-RF, which is experimentally difficult to distinguish from the 80-bp T-RF, corresponded to several taxa. Since the *MspI* hydrolysis did not resolve the MB community structure in this soil properly we cloned PCR products and sequenced 13 environmental *pmoA* genes. Reconstruction of the

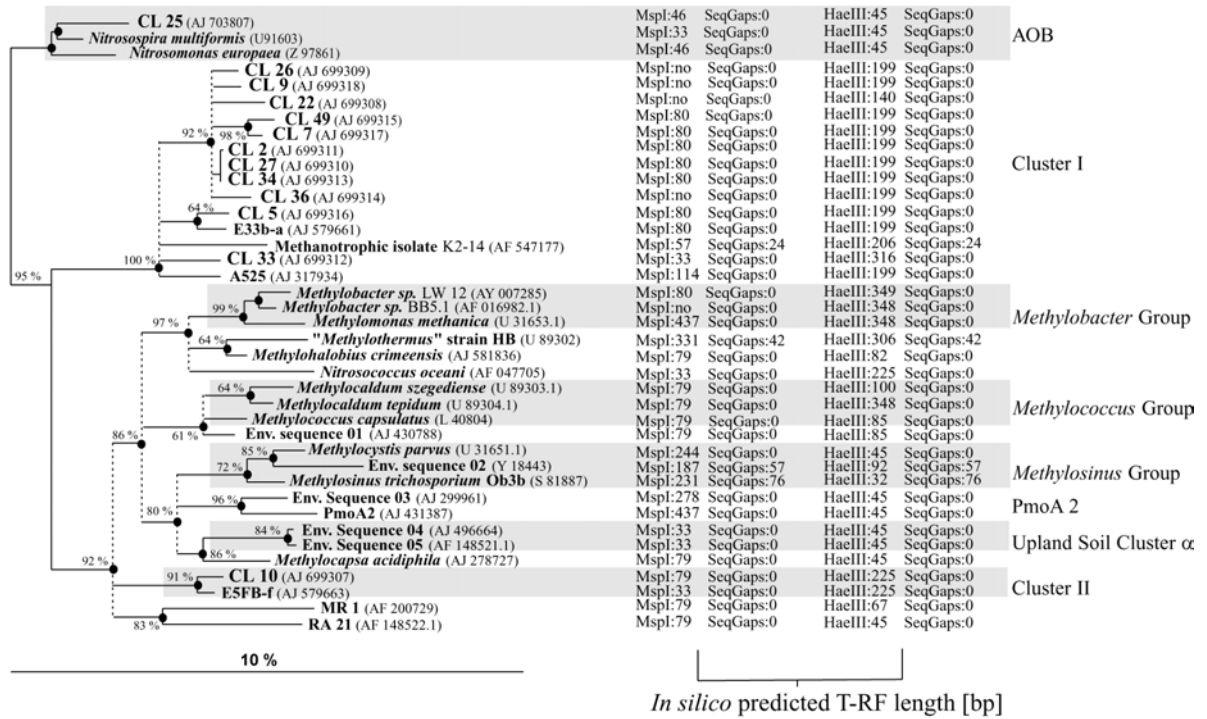
gene phylogeny showed that the majority of sequences (11) were affiliated with Cluster I sequences (Fig. 1A). Only two clones clustered with *amoA* genes (clone 25) and environmental Cluster II sequence E5FB-f (clone 10 (9)), respectively. The lack of clones affiliated with any other group of MB suggested a dominance of Cluster I *pmoA* sequences in this pH-neutral temperate forest soil.

Figure 1: (next page)

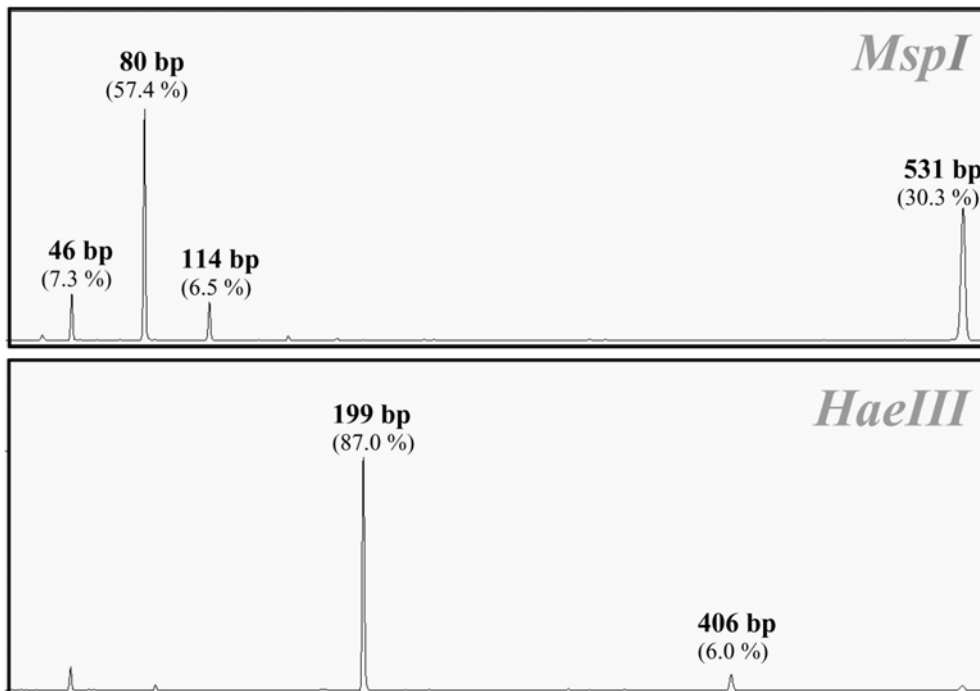
A: Phylogeny of *pmoA* sequences and corresponding T-RFs predicted by *in silico* analysis through TRF-CUT. The tree is slightly modified from the ARB display. The tree was calculated from 127 deduced amino acid positions with the Tree-Puzzle, Maximum likelihood, and Neighbor-joining methods. 10000 puzzling steps were performed and probability estimates ($\geq 60\%$) are given as percentage values at nodes. Numbers in parenthesis indicate accession numbers. Thickened nodes were common to both Maximum likelihood and Neighbor-joining analyses. Dashed vertical lines (polytomies) indicate unresolved branching, which were inferred from conflicting results of the phylogenetic analyses and low ($\leq 60\%$) support values of Tree-Puzzle. The scale bar indicates 10% sequence divergence. T-RF lengths after *MspI* and *HaeIII* *in silico* hydrolysis are given by numbers; no, no restriction site within analysed *pmoA* fragment. SeqGaps 0, no gap; number indicates a gap between priming site and start of the actual sequence.

B: T-RFs retrieved from a pH-neutral temperate forest soil by T-RFLP using *MspI* and *HaeIII*. Numbers indicate lengths of T-RFs (in basepairs) represented by *pmoA* clones; no: no restriction site within the *pmoA* fragment analysed. Numbers in parenthesis are the relative contribution of each T-RF to the total fluorescence of the sample. X axis, size of T-RFs in basepairs; y axis, relative fluorescence units.

A



B



Based on these findings we applied TRF-CUT to test various restriction endonucleases, in order to select one that allowed the discrimination of Cluster I sequence types from all other *pmoA/amoA* genes. *In silico* digestion with *HaeIII* revealed a 199-bp fragment unique to the majority of Cluster I sequences (Fig. 1A). For two Cluster I sequences TRF-CUT indicated equally unique T-RFs of 140-bp and 316-bp sizes, respectively. Subsequent T-RFLP analysis of *pmoA* amplified from the soil DNA extract indeed showed a predominance of a 199-bp T-RF when *HaeIII* was used for digestion (87.0 %), thus confirming the predicted result. An additional 406-bp T-RF corresponded exclusively to the Pseudo-T-RF of Cluster I sequences (3). Summing up the relative abundance of these two peaks, Cluster I MB comprised 93.0 % of the detected MB community. Peaks of the predicted sizes of 140-bp and 316-bp for Cluster I, of 225-bp for Cluster II, and of 45-bp for AOB-affiliated clones, were all relatively small (< 1.0 %) or undetected, indicating that organisms with these genotypes were only minor constituents of the methanotrophic guild in this forest soil.

In summary, application of TRF CUT prior to T-RFLP analysis revealed Cluster I as the predominant MB in this pH-neutral methane-consuming soil, and indicated that the distribution of this novel gene cluster is more widespread than previously known. In this study a clone library was necessary because of the limited availability of *pmoA* sequences from different habitats. However, with the prerequisite of an extensive dataset, *in silico* predictions of multiple enzyme digests should allow assignment of TRFs to potentially corresponding organisms without extensive prior cloning efforts.

We acknowledge Peter F. Dunfield for critically reading this manuscript. This work was financially supported by the German Federal Ministry of Education and Research within the Competence Network Göttingen project Genome Research on Bacteria (GenoMik) and the BIOLOG Biodiversity Program (01LC0021) and by grants of the Max Planck Society (Munich).

References

1. **Avrahami, S., R. Conrad, and G. Braker.** 2002. Effect of soil ammonium concentration on N₂O release and on the community structure of ammonia oxidizers and denitrifiers. *Appl. Environ. Microbiol.* **68**:5685-5692.
2. **Braker, G., H. L. Ayala-del-Rio, A. H. Devol, A. Fesefeldt, and J. M. Tiedje.** 2001. Community structure of denitrifiers, Bacteria, and Archaea along redox gradients in Pacific Northwest marine sediments by terminal restriction fragment length polymorphism analysis of amplified nitrite reductase (*nirS*) and 16S rRNA genes. *Appl. Environ. Microbiol.* **67**:1893-1901.
3. **Egert, M. and M. W. Friedrich.** 2003. Formation of pseudo-terminal restriction fragments, a PCR-related bias affecting terminal restriction fragment length polymorphism analysis of microbial community structure. *Appl. Environ. Microbiol.* **5**:2555-2562.
4. **Henckel, T., U. Jäkel, S. Schnell, and R. Conrad.** 2000. Molecular analysis of novel methanotrophic communities in forest soil that oxidize atmospheric methane. *Appl. Environ. Microbiol.* **66**:1801-1808.
5. **Holmes, A. J., A. Costello, M. E. Lidstrom, and J. C. Murrell.** 1995. Evidence that particulate methane monooxygenase and ammonia monooxygenase may be evolutionary related. *FEMS Microbiol. Lett.* **132**:203-208.
6. **Horz, H.-P., M. T. Yimga, and W. Liesack.** 2001. Detection of methanotroph diversity on roots of submerged rice plants by molecular retrieval of *pmoA*-based terminal restriction fragment length polymorphism profiling. *Appl. Environ. Microbiol.* **67**:4177-4185.
7. **Jensen, S., A. J. Holmes, R. A. Olsen, and J. C. Murrell.** 2000. Detection of methane oxidizing bacteria in forest soil by monooxygenase PCR amplification. *Microbiol. Ecol.* **39**:282-289.
8. **Kent, A. D., D. J. Smith, B. J. Benson, and E. W. Triplett.** 2003. Web-based phylogenetic assignment tool for analysis of terminal restriction fragment length polymorphism profiles of microbial communities. *Appl. Environ. Microbiol.* **69**:6768-6776. [<http://trflp.limnology.wisc.edu>].
9. **Knief, C., A. Lipski, and P. F. Dunfield.** 2003. Diversity and activity of methanotrophic bacteria in different upland soils. *Appl. Environ. Microbiol.* **69**:6703-6714.
10. **Kolb, S., C. Knief, S. Stubner, and R. Conrad.** 2003. Quantitative detection of methanotrophs in soil by novel *pmoA*-targeted real-time PCR assays. *Appl. Environ. Microbiol.* **69**:2423-2439.
11. **Ludwig, W., O. Strunk, R. Westram, L. Richter, H. Meier, K. Yadhu, A. Buchner, T. Lai, S. Steppi, G. Jobb, W. Forster, I. Brettske, S. Gerber, A.**

- W. Ginhart, O. Gross, S. Grumann, S. Hermann, R. Jost, A. König, T. Liss, R. Lussmann, M. May, B. Nonhoff, B. Reichel, R. Strehlow, A. Stamatakis, N. Stuckmann, A. Vilbig, M. Lenke, T. Ludwig, A. Bode, and K.-H. Schleifer.** 2004. ARB: a software environment for sequence data. *Nucl. Acids Res.* **32**:1363-1371. [www.arb-home.de].
12. **Marsh, T. L., P. R. Saxman, J. R. Cole, and J. M. Tiedje.** 2000. Terminal restriction fragment length polymorphism analysis program, a web-based research tool for microbial community analysis. *Appl. Environ. Microbiol.* **66**:3616-3620. [<http://rdp.cme.msu.edu>].
13. **Pacheco-Oliver, M., I. R. McDonald, D. Groleau, J. C. Murrell, and C. B. Miguez.** 2002. Detection of methanotrophs with highly divergent *pmoA* genes from Arctic soils. *FEMS Microbiol. Lett.* **209**:313-319.
14. **Yeager, C. M., J. L. Kornosky, D. C. Housman, E. E. Grote, J. Belnap, and C. R. Kuske.** 2004. Diazotrophic community structure and function in two successional stages of biological soil crusts from the Colorado Plateau and Chihuahuan Desert. *Appl. Environ. Microbiol.* **70**:973-983.

4. Diskussion

4.1 Untersuchungen zu Funktion und Verbreitung der pMMO-2

Eine ungewöhnliche paraloge Kopie des *pmoA*-Gens.

Das Vorkommen zweier stark von einander abweichender Kopien des Gens *pmoA* in einem Genom wurde erstmals von Dunfield und Mitarbeitern nachgewiesen (Dunfield et al., 2002). Zwar war das Vorhandensein von bis zu drei Kopien der für die pMMO kodierenden Gene in einem Genom sowohl für methanotrophe *Alpha*- wie *Gammaproteobacteria* dargestellt worden (Stolyar et al., 1999; Stolyar et al., 2001b), alle analysierten Kopien wiesen jedoch sehr hohe Sequenzähnlichkeiten zueinander auf. Die Existenz der neuartigen *pmoA*-Genkopien (novel *pmoA*, *pmoA2*) wurde zunächst hauptsächlich in Bezug auf die damit verbundene Komplikation *pmoA*-basierter phylogenetischer und populationsökologischer Analysen diskutiert (Dunfield et al., 2002). Beispielsweise gruppierte die zuvor als *pmoA* eines Vertreters einer unkultivierten methanotrophen Linie charakterisierte Sequenz M84P3 (Horz et al., 2001) zusammen mit den zwei von Dunfield und Mitarbeitern aus Reinkulturen gewonnenen *pmoA2*-Sequenzen. Im Zuge der hier diskutierten Arbeit wurde zunächst die Verbreitung von *pmoA2* innerhalb der MB sowie die Phylogenie dieses Gens charakterisiert. Anschließend wurden mittels der Klonierung großer genomischer Fragmente die Operonstrukturen sowohl der konventionellen *pmoA1* als auch der *pmoA2* des Referenzorganismus *Methylocystis* sp. Stamm SC2 untersucht. Dabei konnte der Nachweis erbracht werden, daß *pmoA2* Teil eines vollständigen, als Einheit transkribierten Operons ist. Trotz deutlicher Unterschiede auf DNA-Ebene ergab die Analyse der aus den Gensequenzen abgeleiteten Polypeptide, insbesondere der *in silico* vorhergesagten Sekundärstrukturen, eine grosse Ähnlichkeit von pMMO-1 und pMMO-2.

4.1.1 Untersuchung zur Verbreitung und Phylogenie der *pmoA2*

Mit dem Ziel die Verbreitung des Gens *pmoA2* innerhalb der bekannten MB zu charakterisieren, wurden während dieser Arbeit eine Vielzahl von Typ I, Typ II und Typ X MB untersucht. Der Nachweis von *pmoA2* mittels eines neu entwickelten, für

pmoA2 spezifischen PCR-Assays gelang bei 19 von 27 untersuchten Typ II MB der Gattungen *Methylosinus* und *Methylocystis*. Das Gen wurde jedoch bei keinem untersuchten Typ I / Typ X MB gefunden. Interessanterweise konnte das Gen bei verschiedenen Typ II MB ebenfalls nicht nachgewiesen werden, obwohl es bei nahe verwandten Organismen detektiert wurde. Die mangelnde Nachweisbarkeit ist auf ein Fehlen des Gens in den Genomen der entsprechenden Organismen zurückzuführen, und nicht etwa auf abweichende Basenabfolgen in den während der PCR als Primerbindestellen genutzten Bereichen des Gens. Dies wurde durch „Southern“-Hybridisierungsexperimente mit einer *pmoA2*-spezifischen Sonde bestätigt. In dieser offensichtlich nicht mit der Phylogenie korrelierten Verteilung erinnert *pmoA2* an die für die lösliche Methan-Monooxygenase kodierenden Gene *mmoXYBZC* (Hanson and Hanson, 1996; Heyer et al., 2002b). Die phylogenetische Verrechnung der erhobenen Sequenz-Daten im Vergleich zu Sequenzen des *pmoA1*-Gens zeigt anschaulich die hohe Ähnlichkeit aller partiellen *pmoA2*-Sequenzen. Unabhängig von der phylogenetischen Position der zugehörigen Organismen gruppieren alle *pmoA2*-Sequenzen in einem monophyletischen Cluster. Dabei gruppieren die Linien der *pmoA2*- und der *pmoA1*-Sequenzen von Typ II MB bei allen Berechnungsmethoden von einander getrennt, bilden aber einen gemeinsamen Entwicklungsast, der deutlich abgegrenzt von den *pmoA1*-Sequenzen der Typ I MB verzweigt.

Die Tatsache, daß Genome physiologisch und taxonomisch hochähnlicher Organismen das Gen *pmoA2* enthalten oder auch nicht enthalten, wirft Fragen bezüglich einer möglichen Funktion des Genprodukts wie auch des Entstehungs- und Verbreitungs-Mechanismus dieses Gens auf. Der erste Aspekt wird weiter unten diskutiert (vgl. 4.1.6). Zur Erklärung des Verbreitungsmusters kommen zwei einander nicht ausschließende Mechanismen in Frage: Häufiger Verlust der Genkopie aus dem Genom oder Erwerb durch horizontalen Gentransfer.

Das relative Verhältnis von nicht-synonymen zu synonymen akzeptierten Mutationen gibt Aufschluß darüber, ob ein selektiver Druck auf eine Sequenz einwirkt. Anhand der oben beschriebenen Sequenzen konnte abgeleitet werden, daß nach der Aufspaltung in *pmoA2* und *pmoA1* zunächst ein diversifizierender Druck zu einer starken Veränderung des *pmoA2*-Vorläufer-Gens geführt hat. Im Anschluß setzte eine konservierende Selektion ein, die zeitlich vermutlich mit der Auftrennung in verschiedene methanotrophe Linien innerhalb der Typ II MB einherging (vgl. 3.1.1). Hypothesen zur Entstehung des *pmoA2*-Verbreitungsmusters müssen diesen

Befund erklären können. Eine abschließende Aufklärung des Verbreitungs-Mechanismus der *pmoA2*-Gene ist auf Basis der vorhandenen Daten und Kenntnisse nicht möglich. Allerdings können anhand der hier beschriebenen Teilsequenzen des *pmoA2*-Gens verschiedene Szenarien diskutiert werden:

i) Das gemeinsame Verzweigen von *pmoA2*-Sequenzen und Typ II MB *pmoA1*-Sequenzen in phylogenetischen Analysen sowie die Tatsache, daß *pmoA2* in keinem Typ I / Typ X MB nachgewiesen werden konnte, deuten auf die Entstehung von *pmoA2* durch eine Genduplikation in einem gemeinsamen Vorfahren der Typ II MB (*Alphaproteobacteria*) nach deren Differenzierung von den Typ I / Typ X MB (*Gammaproteobacteria*) hin. Ein Entstehen vor dieser Aufspaltung wäre lediglich durch den raschen Verlust des Genduplikats, sowie einen starken diversifizierenden Druck auf die konventionelle PmoA im hypothetischen Vorläufer der *Gammaproteobacteria* erklärbar. Das beobachtete Verteilungsmuster innerhalb der Typ II MB müßte dementsprechend, eine dychotome Phylogenie vorausgesetzt, durch häufige Gen-Deletionen nach (oder während) der Aufspaltung des gemeinsamen Vorläufers der Typ II MB in die Gattungen *Methylosinus* und *Methylocystis* erklärt werden. Zeitlich vor dieser Separation muß eine verstärkte Evolution der pMMO-2-Vorläufer-Sequenzen postuliert werden: die hohe Ähnlichkeit aller bekannten *pmoA2*-Sequenzen auch auf DNA-Ebene sowie die konservierten Abweichungen von *pmoA1*-Sequenzen dürften kaum durch eine konvergente Entwicklung nach Aufspaltung in die genannten Gattungen erklärbar sein.

ii) Ein alternativer, jedoch eher unwahrscheinlicher Erklärungsansatz wäre der Folgende: die Gen-Duplikation fand nach der Auftrennung der Typ II MB in die Gattungen *Methylosinus* und *Methylocystis* in einem frühen Vertreter einer der beiden Gattungen statt. Auch hier muß ein starker Selektionsdruck über einen kurzen evolutionären Zeitraum postuliert werden, in dem sich die Vorläufer-Sequenz der rezenten *pmoA2*-Gene im Vergleich zu den stark konservierten *pmoA1*-Sequenzen veränderte. Das beobachtete Verteilungsmuster muß nun allerdings durch häufigen horizontalen Gentransfer zu anderen Typ II MB erklärt werden. Die Tatsache, daß dieser Austausch anscheinend nur zwischen Typ II MB stattgefunden hätte, spricht allerdings gegen diese Hypothese. Andererseits ist der häufige Verlust eines unter konservierendem selektivem Druck stehenden, und somit möglicherweise funktionell wichtigen Gens während der Evolution im Vergleich zur Aquisition eines solchen Gens durch horizontalen Gentransfer als eher unwahrscheinlich zu bewerten.

Natürlich kann auch das parallele Einwirken beider Mechanismen nicht ausgeschlossen werden.

4.1.2 Vergleichende Analyse der Operons *pmoCAB1* und *pmoCAB2* von *Methylocystis* sp. Stamm SC2

Die weite Verbreitung von *pmoA2* in Typ II MB, der durch die Mutations-Analyse nachgewiesene konservierende Selektionsdruck sowie der mittels RT-PCR geführte Nachweis zumindest einer schwachen Expression des *pmoA2*-Gens in *Methylocystis* sp. Stamm SC2 unter Standard-Wachstumsbedingungen wurden in ihrer Gesamtheit als Hinweis auf eine aktive Funktion der neuartigen PmoA2 gedeutet. Ein weiteres Ziel dieser Arbeit bestand deshalb darin, diese Hypothese durch eine Analyse der flankierenden genomischen Region des *pmoA2*-Gens zu überprüfen. Dabei wurde anhand des Model-Organismus *Methylocystis* sp. Stamm SC2 untersucht, ob die von konventionellen *pmoA*-Genen bekannte Operonstruktur (Gilbert et al., 2000; Stolýar et al., 1999) auch bei *pmoA2* vorhanden ist, insbesondere ob ein vollständiges, von *pmoC2* und *pmoB2* flankiertes Gen existiert. Durch Vergleich mit der Struktur des konventionellen Operons, durch phylogenetische Analysen der Primär- und vergleichende Analysen der Sekundärstruktur der abgeleiteten Aminosäure-Sequenzen konnte nachgewiesen werden, daß es sich bei pMMO-2 um ein ebenfalls Membran-assoziiertes Enzym handelt, das viele der von pMMO-1 bekannten Eigenschaften aufweist. In einer Fortführung der Expressionsstudien von Tchawa Yimga und Mitarbeitern (2003) wurden die Transkriptionsstartpunkte mittels der RACE-Technik ("Rapid Amplification of cDNA Ends") experimentell bestimmt, was die anschließende Identifikation und den Vergleich beider Promoter-Regionen ermöglichte.

Erstellung einer BAC-Bibliothek aus genomischer DNA von *Methylocystis* sp. Stamm SC2.

Autoren verschiedener Studien hatten von Schwierigkeiten berichtet, komplette pMMO-Operons in *E. coli* zu klonieren (Gilbert et al., 2000; Semrau et al., 1995b). Als mögliche Ursache wurden toxische Effekte des Genprodukts diskutiert. Allen Studien war gemein, daß mit Vektoren gearbeitet wurde, die in hoher Kopienzahl ("High Copy Vector") mit mehreren 100 Kopien pro Zelle vorliegen. Um die

Problematik zu umgehen, wurde die Klonierung mittels BACs, die nur mit ein bis zwei Kopien pro Zelle vorliegen, durchgeführt. Eine ca. 1000 Klone umfassende BAC-Bibliothek mit einer durchschnittlichen Insertgröße von 80 kb wurde mit spezifischen PCR-Assays sowohl hinsichtlich *pmoA1* als auch hinsichtlich *pmoA2* gescreent, jeweils ein testpositiver Klon wurde anschließend sequenziert.

Vergleich der Operonstrukturen von pMMO-1 und pMMO-2.

Die stromaufwärts bzw -abwärts des *pmoA*-Gens identifizierten offenen Leserahmen (ORFs) konnten eindeutig als *pmoC* bzw. *pmoB* charakterisiert werden. Die Längen dieser Gene liegen in einem Bereich, der zuvor bereits für die entsprechenden Gene zweier Typ II und eines Typ X MB berichtet worden ist (vgl. 3.1.2, Tab. 2). Der paarweise Vergleich zueinander homologer Gene ergab prozentuale Sequenz-Ähnlichkeiten derselben Größenordnung, die zuvor für das zentrale Fragment der *pmoA* bestimmt worden war. Daraus kann abgeleitet werden, daß alle drei Gene eine vergleichbare phylogenetische Entwicklung durchlaufen haben. Phylogenetische Analysen der Aminosäuresequenzen aller drei abgeleiteten Polypeptide bestätigen dies (vgl. 3.1.2). Im Gegensatz zu den kodierenden Bereichen wiesen die intergenischen Regionen der zwei verglichenen Operons keine signifikanten Übereinstimmungen auf. Die Zuordnung homologer Nukleotide (Alignment) dieser Bereiche war nicht möglich. Dieses Resultat unterstützt die Annahme, daß die Gene der pMMO-2 einem konservierenden evolutionären Druck unterworfen sind, der eine Akkumulation von Mutationen verhinderte, wie sie für die während des gleichen Zeitraums evolvierten intergenischen Regionen festgestellt wurde. Stromaufwärts der *pmoCAB*-Gene beider Operons wurden Ribosomen-Bindestellen („Shine-Dalgarno-Sequenzen“) identifiziert. Die Konserviertheit dieser innerhalb der hochvariablen intergenischen Regionen gelegenen Motive deutet darauf hin, daß die Shine-Dalgarno-Sequenzen noch funktionell bedeutsam sind. Dieser Befund bereitet zusätzliche Evidenz, daß die Gene der pMMO-2 nicht nur transkribiert sondern auch translatiert werden. Beide Gencluster werden stromaufwärts durch Promoter-Bereiche (vgl. 4.1.5), stromabwärts durch Faktor-unabhängige Transkriptionsterminatoren flankiert. Alle notwendigen Elemente für Transkription und Translation konnten nachgewiesen werden. Somit wurde erstmals gezeigt, daß die Gene *pmoC2*, *pmoA2* und *pmoB2* in einem Operon organisiert sind, dessen Struktur mit der Struktur konventioneller pMMO-Operons übereinstimmt.

Analyse der abgeleiteten Aminosäuresequenzen von pMMO-1 und pMMO-2

Während die *Identitätswerte* der abgeleiteten Aminosäuresequenzen von pMMO-1 und pMMO-2 durchweg sogar unter den Identitätswerten der entsprechenden Gene liegen, wurden deutlich höhere *Ähnlichkeitswerte* festgestellt. Zur Berechnung der Identität wird der prozentuale Anteil identischer Aminosäuren zweier Sequenzen bestimmt. Die Ähnlichkeit ist definiert als relativer Anteil *funktionell ähnlicher* Aminosäuren. Es werden also alle Positionen eines Alignments gezählt, an denen die zu vergleichenden Polypeptide Aminosäurereste mit vergleichbaren physikochemischen bzw. strukturellen Eigenschaften aufweisen. Hohe Ähnlichkeitswerte deuten auf eine vergleichbare Funktion zweier Proteine hin. Wesentlich deutlicher konnte die mögliche funktionelle Entsprechung von pMMO-1 und pMMO-2 durch die vergleichende Ableitung der Sekundärstrukturen beider Proteine dargestellt werden. Anhand verschiedener, auf statistischen Modellen ("Hidden Markov Modelle") wie auch auf der Hydrophobizität einzelner Aminosäuren basierender Analysemethoden (vgl. 2.1.3) konnten die wahrscheinlichen Sekundärstrukturen aller Polypeptide, insbesondere eine große Zahl transmembraner Helices, vorhergesagt werden. Die Sekundärstruktur von pMMO-2 stimmt hinsichtlich Anzahl und Lage der Transmembran-Regionen sowohl mit den drei bereits bekannten Sekundärstrukturen konventioneller pMMOs (Gilbert et al., 2000; Stolyar et al., 1999) als auch im direkten Vergleich mit der Struktur der pMMO-1 von *Methylocystis* sp. Stamm SC2 stark überein (vgl. 3.1.2, Abb. 3). Strukturelle Ähnlichkeiten gelten allgemein als starkes Indiz für ähnliche Funktionen. Das trotz der auf Ebene der Primärstrukturen von DNA und Aminosäure beobachteten prozentualen Distanzen eine große strukturelle Übereinstimmung beider Proteine besteht, muß als beträchtliche Evidenz dafür gewertet werden, daß pMMO-2 einem konservierenden evolutionären Druck unterworfen ist. Die untersuchten Gene stellen somit keine Pseudogene dar, sondern codieren für ein Enzym, das die Reaktion einer membranassoziierten Monooxygenase ausführen kann.

Dieser Befund wird unterstützt durch eine vergleichende Analyse hochkonservierter Aminosäurereste, die bei nahezu allen bekannten pMMO- und AMO-Sequenzen vorhanden sind. Dazu wurde auf die Arbeit von Tikhvatullin und Mitarbeitern (2000) zurückgegriffen, in der insgesamt 203 hochkonservierte Reste identifiziert wurden. Die Mehrzahl dieser Aminosäurereste (190) ist auch an den

homologen Positionen der pMMO-2 von *Methylocystis* sp. Stamm SC2 vorhanden. Die Abweichung an 13 Positionen von einer Konsensussequenz, die sowohl für die bekannten Sequenzen der Typ I und Typ II MB als auch für Sequenzen der homologen AMO charakteristisch ist, erscheint um so erstaunlicher, da pMMO-2 phylogenetisch zusammen mit pMMO-1 Sequenzen der Typ II MB gruppiert, während die ebenfalls der Konsensussequenz zugrundeliegenden AMO-Sequenzen und pMMO-1 Sequenzen der Typ I MB deutlich getrennt verzweigen. Der Befund steht in Übereinstimmung mit der unter 4.1.1 diskutierten evolutionären Entwicklung von pMMO-2. Er gibt zu Vermutungen Anlaß, daß das Enzym eine der konventionellen pMMO zwar ähnliche, jedoch abweichende Funktion haben und beispielsweise eine Anpassung an bestimmte Umweltbedingungen darstellen könnte. Obwohl der hohe Grad an Konserviertheit darauf hindeutet, daß die von Tikhvatullin und Mitarbeitern beschriebenen Aminosäure-Reste bei der enzymatischen Reaktion eine wichtige Rolle spielen, muß nicht von einem Verlust der katalytischen Aktivität der pMMO-2 ausgegangen werden, da 12 der 13 von der Konsensussequenz abweichenden Reste der gleichen funktionellen Aminosäuren-Gruppe wie ihre konservierten homologen Gegenstücke angehören. Der Ähnlichkeits-Wert der 203 verglichenen Reste von *pmoCAB2* und Konsensussequenz liegt somit über 99%. Tikhvatullin und Mitarbeiter analysierten anhand von Struktur-Daten (Mössbauer Spektroskopie, ESR u.a.) das Potential der hochkonservierten Aminosäurereste, als Bindestellen der in der pMMO enthaltenen Metallatome zu fungieren. Solche Reste können bei der Ausbildung des aktiven Zentrums eine essentielle Rolle spielen. Von 39 konservierten, potentiell Liganden-bindenden Aminosäuren sind 38 auch in der pMMO-2 von *Methylocystis* sp. Stamm SC2 vorhanden. Die einzige Ausnahme (Y26->F in *pmoA*) ist in der Mehrzahl der vorgeschlagenen Modelle nicht an der Formation des katalytischen Zentrums beteiligt. Darüber hinaus sind Tyrosin (Y) und Phenylalanin (F) durch ähnliche strukturelle und chemische Eigenschaften charakterisiert. Deshalb muß der Austausch nicht zwangsläufig einen Einfluß auf die katalytische Aktivität der pMMO-2 gehabt haben.

Die Ableitung der Funktion einer pMMO/AMO nur aufgrund der phylogenetischen Stellung ihrer Sequenzen ist ohne Kenntnis der physiologischen Eigenschaften des Organismus nicht möglich. Beispielsweise sind die AmoA-Sequenzen nitrifizierender *Gammaproteobacteria* phylogenetisch näher mit PmoA-Sequenzen von *Gammaproteobacteria* verwandt, als mit den AmoA-Sequenzen der

Betaproteobacteria. Um die pMMO-2 funktionell gegen die AMO abzugrenzen, wurde eine Signaturanalyse durchgeführt. Der Vergleich von 919 AmoA-, 315 PmoA1- und 40 PmoA2-Sequenzen ermöglichte die Bestimmung von 18 Positionen, an denen AmoA und PmoA1 unterschiedliche Signaturen (konservierte Reste) aufweisen, die in der Gesamtheit eine Differenzierung beider Enzymfunktionen erlauben. Da die PmoA2-Konsensussequenz hohe Übereinstimmung mit den Signaturen der PmoA1 aufweist (vgl. 3.1.2, Tab. 3), wurde pMMO-2, entsprechend der bisherigen Annahmen, die Funktion einer Methan-Monooxygenase zugewiesen.

Bestimmung der Transkriptionstartpunkte mittels RACE

Durch reverse Transkription der mRNA von *Methylocystis* sp. Stamm SC2 konnte spezifische cDNA für beide Operonkopien generiert werden. Polyadenylierung der 3'-Enden ermöglichte die Amplifikation des 3'-terminalen Bereichs der cDNA durch Kombination Operon-spezifischer Primer mit Oligo-T-Primern. Durch anschließende Sequenzierung der Amplifikate konnte das zur ersten Base der mRNA korrespondierende 3'-Ende der cDNA basengenau festgestellt werden (vgl. 2.10.7). Die Ergebnisse wurden anschließend mittels spezifischer PCR-Analysen überprüft und bestätigt. Stromaufwärts der Transkriptions-Startpunkte konnten sowohl für *pmoCAB1* wie auch für *pmoCAB2* Nukleotid-Sequenzen identifiziert werden, die hohe Ähnlichkeit zu der Konsensussequenz der -35 und -10 Elemente von *E. coli*- σ^{70} -Promotoren aufweisen. Der Bereich wurde dementsprechend als Promoter annotiert. Ein Vergleich mit den von Gilbert und Mitarbeitern bestimmten Promotoren der *pmo*-Operons von *M. trichosporium* OB3b und *Methylocystis* sp. M zeigte (Gilbert et al., 2000), daß die -35 und -10 Elemente konventioneller *pmo*-Operons untereinander eine hohe Sequenzähnlichkeit aufweisen. Der *pmo2*-Promoter weist deutliche Unterschiede zu diesen Sequenzen auf. Wenngleich andererseits die Ähnlichkeit aller Promotoren zur *E.coli*-Konsensussequenz vergleichbar ist, könnte dies ein Indiz dafür sein, daß es sich bei dem *pmo2*-Promoter um einen, im Vergleich zu *pmo1*-Promotoren, „schwachen“ Promoter handelt. Der Befund steht in Übereinstimmung mit der im Vergleich zur pMMO-1-Transkription deutlich schwächeren Transkription von pMMO-2 unter Standardwachstumsbedingungen (vgl. 4.1.6). Eine gezielte Regulation als Antwort auf veränderte Umweltbedingungen

dürfte eher durch zusätzliche Kontrollelemente erfolgen, als durch Verwendung eines alternativen σ -Faktors.

Mögliche Funktion der pMMO-2

Zusammengefaßt deuten die diskutierten Ergebnisse, insbesondere der durch verschiedene Analysen erwiesene, auf pMMO-2 wirkende Selektionsdruck, die experimentelle Darstellung der Transkription (RACE-Analysen) sowie die weite Verbreitung der *pmoA2* innerhalb der Typ II MB auf eine aktive Funktion des Enzyms hin. Die Tatsache, daß pMMO-2 bei vielen MB nicht vorhanden ist, steht nicht im Widerspruch zu dieser Hypothese, da ein ähnliches Verbreitungsmuster beispielsweise auch für die funktionell wichtige sMMO bekannt ist. Die sMMO wird alternativ zur pMMO nur bei Kupferkonzentrationen unter 0,2 μ M exprimiert.

Insbesondere die Abweichungen der Sequenzen von PmoCAB von der ansonsten für pMMO und AMO stark charakteristischen Konsensussequenz deuten auf eine mögliche "Spezialisierung" der pMMO-2 hin. Allerdings können über die genaue Funktion der pMMO-2, insbesondere über die Unterschiede zur konventionellen pMMO-1, nur Vermutungen angestellt werden. Ein erster Versuch zur Bestimmung der funktionellen Besonderheiten wurde mit einer im Rahmen dieser Promotion angeleiteten Diplomarbeit von Viola Paulus unternommen (Paulus, 2004). Darin wurde mittels spezifischer Amplifikation von cDNA das relative Verhältnis der Transkripte beider *pmo*-Operons unter verschiedenen Wachstumsbedingungen getestet. Ein so detektierter Anstieg des *pmoCAB2*-Transkriptions-Niveaus bei Kultivierung von *Methylocystis* sp. Stamm SC2 unter verminderter Sauerstoff-Konzentration (1% O₂ statt 21% unter Standardbedingungen) konnte jedoch durch direkten Nachweis der mRNA mittels Northern-Hybridisierung nicht bestätigt werden. Dennoch werden die Ergebnisse als erster Hinweis auf eine verstärkte Expression der pMMO-2 unter *in situ* Bedingungen gewertet. Die in natürlichen Habitaten gemessenen O₂-Konzentrationen der oxisch/anoxischen Grenzschichten liegen in der Regel deutlich unter 1 %, konnten im Labor jedoch nicht reproduziert werden.

Zwar konnte die Expression von *pmoCAB2* unter allen getesteten Wachstumsbedingungen nachgewiesen werden, von Paulus bzw. im Rahmen der vorliegenden Arbeit durchgeführte semiquantitative Analysen ergaben jedoch stets ein

deutlich geringeres Transkriptionsniveau als für *pmoCAB1*. Ob pMMO-2 eine von der pMMO-1 (leicht) abweichende Funktion aufweist und als Isoenzym eine Anpassung an bestimmte Umweltfaktoren darstellt, kann zum gegenwärtigen Zeitpunkt nicht eindeutig geklärt werden. Die Frage wird frühestens mit dem Abschluß von unlängst begonnenen Studien beantwortet werden können, die die Auswirkungen des gezielten Ausschaltens sowohl der konventionellen als auch der neuartigen Form der pMMO (Analyse von Knockout-Mutanten) untersuchen.

4.2 Erste Einblicke in ein Genom nicht-kultivierter MB

4.2.1 USC α : langgesuchte hoch-affine MB ?

Trotz zahlreicher Bemühungen ist es bislang nicht gelungen, zweifelsfrei diejenigen MB zu bestimmen, die die Oxidation atmosphärischen Methans in als Methan-Senken charakterisierten Böden katalysieren. Wenngleich nicht ausgeschlossen werden kann, dass Vertreter bereits gut erforschter methanotropher Linien zur fakultativ hoch-affinen Methan-Oxidation befähigt sind, sprechen andererseits viele Befunde für die Oxidation atmosphärischen Methans durch bislang nicht kultivierte MB. So wurden beispielsweise in verschiedenen, als Methan-Senken fungierenden Böden predominant Sequenzen der Gruppen USC α und USC γ nachgewiesen (Holmes et al., 1999; Kolb et al., 2003). Alle Bemühungen, Vertreter dieser Linien zu kultivieren oder anzureichern, scheiterten bislang (Knief, 2004).

Ziel dieses Teils der vorliegenden Arbeit war es deshalb, mittels der Metagenomik einen ersten Einblick in das Genom eines Vertreters des USC α zu gewinnen. Als Referenz wurde ein 100 kb-Fragment des acidophilen MB *M. acidiphila* sequenziert. Bioinformatische Analysen der erzielten Fragmente erlaubten die zweifelsfreie phylogenetische Zuordnung von USC α zu den *Alphaproteobacteria*.

Ein weiteres Ziel der nachfolgend diskutierten Versuche war die vergleichende Analyse der pMMO-Operons von USC α und *M. acidiphila* im Kontext bekannter homologer Operons sowie die Analyse der daraus abgeleiteten Polypeptid-Sequenzen. Wenngleich keine aussergewöhnlichen Merkmale der pMMO des USC α nachgewiesen wurden, die als molekulare Grundlage einer erhöhten Affinität des Enzyms zu Methan interpretiert werden könnten, konnte andererseits die hohe Konserviertheit der Struktur der pMMO auch bei nicht den klassischen MB (Typ I und Typ II MB) zugehörigen Methanotrophen gezeigt werden.

4.2.2 Erstellung und Screening von Large-Insert-Bibliotheken

Im Vorfeld der Versuche wurden Methoden der Extraktion hochmolekularer DNA aus Reinkulturen sowie aus Böden etabliert und für die Erstellung von Large-Insert-

Genom-Bibliotheken eingesetzt. Für den Aufbau der Metagenom-Bibliothek wurde hochmolekulare DNA aus Marburger Waldboden extrahiert. Um die Degradation der DNA durch Scherung zu minimieren, mußte auf physikalische Lyse- und Extraktions-Methoden verzichtet werden. Um andererseits eine unbeabsichtigte Abreicherung der Zielgruppe USC α durch präferentiellen Aufschluß anderer Organismen auszuschließen, wurden verschiedene sanfte Aufschluß-Methoden getestet und mittels eines semi-quantitativen PCR-Ansatzes verfolgt (Daten nicht gezeigt). Die Hauptschwierigkeit bei der Extraktion von DNA aus Boden liegt in der starken Kontamination der DNA durch Huminsäuren. Durch Erweiterung vorhandener Aufreinigungs-Protokolle um zusätzliche Waschschrte konnte eine hinlängliche Reinheit der extrahierten DNA erreicht werden. Die Klonierung der DNA in Fosmide resultierte in einer rund 250.000 Klone umfassenden Metagenom-Bibliothek. Die durchschnittliche Insertgröße von ca. 38-40 kb war durch Verwendung der Fosmid-Technologie vorgegeben. Insgesamt sind ca. 10 Gb an genomischer Information in der Bibliothek enthalten. Die Referenz-Bibliothek des *M. acidiphila*-Genoms wurde mittels der BAC-Technologie erstellt. Die 1054 Klone weisen eine durchschnittliche Insertgröße von 70-80 kb auf, was ca. 20 Genom-Äquivalenten von *M. acidiphila* entspricht.

Mittels eines neu etablierten, *pmoA*-spezifischen, PCR-basierten Hoch-Durchsatz-Screenings (vgl. 2.6.9) war es möglich, zwei Klone zu identifizieren, deren Inserts eindeutig dem USC α zugeordnet werden konnten. Die Sequenzierung der Inserts ergab leider eine 100%ige Übereinstimmung der erzielten Sequenzen auf dem gesamten überlappenden Bereich von ca. 30 kb. Deshalb muß angenommen werden, daß die analysierten Fragmente einem einzigen, oder zwei sehr nahe verwandten Genotypen zuzuordnen sind. Da für das Screening zwei verschiedene PCR-Assays verwendet wurden, die universell für *pmoA*-Gene des USC α bzw. universell für nahezu alle bekannten *pmoA*- und *amoA*-Sequenzen sind, sollte das Erzielen zweier hoch-ähnlicher Fragmente nicht auf ein Screening-Artefakt zurückzuführen sein.

Obwohl der Vergleich von genomischen Fragmenten weniger nah verwandter Vertreter des USC α sicherlich eine breitere Analyse der Gruppe ermöglicht hätte, stellt die Identifikation der USC α -Fragmente dennoch eines der ersten Beispiele für die *gezielte* Anwendung der Metagenom-Strategie für die Analyse von in der Umwelt nur in geringer Abundanz vorhandenen phylogenetischen Linien dar.

4.2.3 Phylogenetische Zuordnung des USC α

Analyse konservierter Gene und Gencluster des USC α -Fragments

Neben dem unter 4.2.4 eingehend diskutierten pMMO-Gencluster (ORFs 45,46,48,50) wurden 28 ORFs identifiziert, die signifikante Ähnlichkeiten zu Genen aus öffentlichen Datenbanken aufwiesen. BLAST-Analysen ergaben für 26 der 28 ORFs höchste lokale Sequenzähnlichkeiten zu Genen von Mitgliedern der *Alphaproteobacteria*, insbesondere von *B. japonicum* und *R. palustris*, die beide zu den *Bradyrhizobiaceae* gehören. Lediglich für ein Gen, ORF16, konnten keine homologen Gene in den öffentlich zugänglichen *Alphaproteobacteria*-Genomen nachgewiesen werden. Die konstant hohe Ähnlichkeit zu Genen der *Rhizobiales*, insbesondere der *Bradyrhizobiaceae* ist ein eindeutiger Hinweis auf den alphaproteobakteriellen Ursprung des USC α . Diese Interpretation wird durch die Analyse konservierter Gencluster klar unterstützt. Als konservierte Gencluster werden hier genomische Abschnitte bezeichnet, die aus mindestens zwei aufeinanderfolgenden Genen bestehen, die in der gleichen Anordnung auch in anderen Genomen nachgewiesen wurden. Die übereinstimmende Existenz homologer Gencluster ist ein deutliches Anzeichen eines gemeinsamen evolutiven Ursprung der verglichenen genomischen Abschnitte. 15 der 28 konservierten ORFs des USC α -Fragments (ohne pMMO-Cluster) sind in insgesamt vier konservierten Genclustern organisiert. Die große Ähnlichkeit des USC α -Fragments zu den Genomen der *Bradyrhizobiaceae* wird dadurch unterstrichen, daß für das USC α -Fragment kein Gencluster nachgewiesen werden konnte, das nicht auch im *B. japonicum*-Genom vorhanden ist. Während für verschiedene andere alphaproteobakterielle Genome Übereinstimmungen mit drei (*R. palustris*) bzw. zwei Clustern gezeigt werden konnten, wurde in nicht-alphaproteobakteriellen Genomen höchstens ein zu einem USC α -Gencluster homologer Bereich gefunden.

SOM-basierte phylogenetische Analyse der USC α -Genom-Signatur

Eine unabhängig Bestätigung der Zuordnung des USC α zu den *Alphaproteobacteria* konnte mittels der XanaMap-Software (www.xanagene.com) gewonnen werden. Während die oben diskutierten Ergebnisse auf Basis der Ähnlichkeit größerer Sequenzabschnitte (Gene bzw. Gencluster) gewonnen wurden, analysiert XanaMap

statistisch die relative Verteilung der Di-, Tri- und Tetranukleotide innerhalb eines genomischen Fragments (vgl. Erläuterungen unter 2.11.10). Diese für einzelne taxonomische Gruppen jeweils charakteristische Genom-Signatur erlaubt die phylogenetische Zuordnung des analysierten Fragments unabhängig von der Identifikation einzelner Gene. Die Methode steht beispielsweise in Analogie zu vergleichenden Analysen der Codon-Usage bzw. des G+C-Gehalts, ist jedoch aufgrund der hohen Spezifität der Genom-Signaturen für individuelle taxonomische Gruppen wesentlich aussagekräftiger. Komparative Signaturanaylsen werden detailliert beispielsweise von Deschavanne und Mitarbeitern (1999), Karlin und Mitarbeitern (1992, 1994) oder Sandberg und Mitarbeitern (2001, 2003) diskutiert. Die Signifikanz der durch die XanaMap-Software erzielbaren Zuordnungen wurde beispielsweise von Kanaya und Mitarbeitern (2001) oder Abe und Mitarbeitern (2003) erörtert.

Nahezu alle analysierten USC α -Subfragmente (vgl. 2.11.10) wurden eindeutig den *Alphaproteobacteria* zugeordnet. Die einzige Ausnahme bildet ein Fragment, das im Grenzbereich zwischen *Alpha*- und *Betaproteobacteria* lokalisiert wurde. Dieses Subfragment enthält ORF16, ein potentielles Gen, das lediglich zu einem hypothetischen Gen von *N. europaea* (*Betaproteobacteria*) signifikante Sequenz-Ähnlichkeiten aufweist. Diese Ergebnisse deuten klar darauf hin, daß ORF16 durch horizontalen Gentransfer aufgenommen wurde und betaproteobakteriellen Ursprungs ist.

Zuordnung des USC α innerhalb der *Alphaproteobacteria*

Wenngleich die diskutierten Ergebnisse unzweifelhaft die Zugehörigkeit des USC α -Fragments zu den *Alphaproteobacteria* belegen, sollte die hohe Ähnlichkeit zu den Genomen der *Bradyrhizobiaceae* nicht überinterpretiert werden. Die für das 100 kb-Fragment von *M. acidiphila* durchgeführten Analysen führten zu vergleichbaren Resultaten. Beispielsweise ergab die BLAST-Analyse der ORFs des *M. acidiphila*-Fragments ebenfalls häufig höchste lokale Sequenz-Übereinstimmungen mit Genen von *B. japonicum*. Darüberhinaus konnten sechs Gencluster identifiziert werden, die ebenfalls im *B. japonicum*-Genom konserviert sind. Auch die Resultate der Signatur-Analyse des *M. acidiphila*-Genom-Fragments erinnert stark an die in der USC α -Analyse erzielten Ergebnisse. Ein direkter Vergleich beispielsweise mit anderen

alphaproteobakteriellen Methanotrophen der *Methylocystaceae* (Typ II MB) ist aufgrund des Mangels an Sequenz-Daten dieser Organismen ebenso wenig möglich, wie ein Vergleich mit Vertretern von *Beijerinkea* spp., den auf Ebene der 16S rRNA-Gene zu *M. acidiphila* nächstverwandten Organismen. Leider wurde neben *pmoCAB* weder im USC α - noch im *M. acidiphila*-Fragment Gene identifiziert, für die Sequenz-Informationen einer der genannten Gruppen vorhanden sind. Auch die Verrechnung aller in GenBank vorhandenen, nicht-redundanten Sequenzdaten von *Methylosinus* spp. und *Methylocystis* spp. im Rahmen der Signatur-Analyse erlaubte keine Zuordnung des USC α auf Gattungsebene.

Die vergleichende Analyse der erzielten Genomabschnitte führte jedoch zur Identifikation von fünf Genen, die auf den analysierten Fragmenten sowohl von USC α wie auch von *M. acidiphila* vorhanden sind. Homologe Kopien aller fünf Gene wurden auch in den Genomen der komplett sequenzierten *Bradyrhizobiaceae* (*B. japonicum* und *R. palustris*) gefunden. Um eine höhere phylogenetische Auflösung zu ermöglichen, insbesondere um die bislang nur auf Basis von *pmoA*-Sequenzen postulierte Verwandtschaft von USC α und *M. acidiphila* zu überprüfen, wurden die Phylogenien von vier der fünf Gene rekonstruiert. Für die durch das fünfte Gen kodierte, putativen Membran-Proteine konnte keine phylogenetische Rekonstruktion vorgenommen werden, da sowohl in dem analysierten USC α -Fragment wie auch in den Genomen der *Bradyrhizobiaceae* zu viele paraloge Gen-Kopien vorhanden sind.

In allen vier phylogenetischen Analysen gruppierten die abgeleiteten Protein-Sequenzen von *M. acidiphila* und USC α innerhalb der *Alphaproteobacteria*. Mit Ausnahme der Phylogenie der putativen CowB-Proteine bilden sie stets eine gemeinsame Linie. Die Analyse der als *cowB* annotierten Gene ergab, daß ORF30 des USC α -Fragments und ORF46 des *M. acidiphila*-Fragments für zwei zueinander zwar homologe, jedoch funktionell nicht äquivalente Proteine kodieren. Beide ORFs verzweigen jeweils mit einer anderen paralogen Gen-Kopie des *Mesorhizobium loti*-Genoms. Die CowB-Phylogenie kann deshalb nicht als Interpretations-Grundlage für die phylogenetische Zuordnung des USC α dienen. In den Stammbäumen der "Staphylococcen-ähnlichen Nuklease", der "putativen Zink-Metallo-Protease" sowie der "putativen Prephenate-Dehydratase" verzweigt der *M. acidiphila*/USC α -Ast stets gemeinsam mit den Genen der *Bradyrhizobiaceae*, was ein weiterer Hinweis auf die hohe Ähnlichkeit dieser Organismen ist.

In ihrer Gesamtheit bestätigen die phylogenetischen Analysen der sowohl für USC α wie auch für *M. acidiphila* bekannten, abgeleiteten Aminosäure-Sequenzen den alphaproteobakteriellen Ursprung des USC α sowie die bereits auf Basis von *pmoA*-Sequenzen postulierte nahe Verwandtschaft zu *M. acidiphila*. Die festgestellten Sequenz-Ähnlichkeiten (Sequenz-Identitäten) liegen im Bereich der Werte, die beispielsweise bei Sequenz-Vergleichen innerhalb der Familie *Bradyrhizobiaceae* festgestellt wurden (vgl. Abb. in 3.1.2). Ähnlichkeitswerte, die für pMMO-Sequenzen berechnet wurden, sind für USC α und *M. acidiphila* konstant niedriger als Referenzwerte die für Typ II MB ermittelt wurden. Wenngleich *M. acidiphila* auf Basis aller vorgestellten Berechnungen klar als nächster kultivierter Verwandter des USC α gelten muß, sind die Sequenz-Ähnlichkeiten somit nicht hoch genug, um die Zugehörigkeit dieser acidophilen MB zu einer Gattung vermuten zu können.

4.2.4 Vergleichende Analyse der pMMO-Operons von USC α und *M. acidiphila*

Vergleich von Primär und Sekundärstrukturen

Bei Typ X und Typ II MB liegen die für pMMO kodierenden Gene in einem Operon in der Reihenfolge *pmoCAB* vor (Gilbert et al., 2000; Ricke et al., 2003; Stolyar et al., 1999). Die Analyse der genomischen Fragmente von USC α und *M. acidiphila* ergab, daß diese Organisation auch bei diesen nicht dem klassischen Typ-System einzugliedernden MB hochkonserviert ist. Die Identitätswerte auf DNA-Ebene sowie die Identitäts- und Ähnlichkeitswerte auf Aminosäure-Ebene sind mit Werten vergleichbar, die zuvor für partielle *pmoA*-Sequenzen berechnet worden waren. Dabei liegen die Ähnlichkeitswerte für PmoB unterhalb, die Werte der PmoA oberhalb des Operon-Durchschnitts. Dieses Muster ist bereits von vergleichbaren Analysen anderer *pmo*-Operons bekannt. Die hohe Konserviertheit von PmoA korrespondiert mit der Tatsache, daß das Polypeptid das aktive Zentrum der pMMO enthält. Der direkte Vergleich der homologen Gene/Polypeptide von *M. acidiphila* und USC α resultierte durchgängig in höheren Sequenz-Übereinstimmungen als Vergleiche mit den entsprechenden Sequenzen anderer MB. Andererseits liegt die durchschnittliche Sequenz-Ähnlichkeit (-Identität) mit 81,5 % (69,4%) deutlich unter den in Vergleichen von Typ II MB verschiedener Gattungen berechneten Werten. Die

durchschnittliche Ähnlichkeit (Identität) der pMMO von *Methylosinus trichosporium* OB3b und *Methylocystis* sp. Stamm M beträgt beispielsweise 92% (83,5%).

Anhand verschiedener, unter 2.1.3 beschriebener Analysemethoden konnten die wahrscheinlichen Sekundärstrukturen aller pMMO-Polypeptide berechnet werden. Die Berechnung der Membran-spannenden Regionen resultierte sowohl für die pMMO von *M. acidiphila* wie auch für die pMMO des USC α in Vorhersagen, die sehr gut mit Berechnung der pMMO von Typ X und Typ II MB übereinstimmen. Die Sekundärstruktur der pMMO ist somit, unabhängig von der phylogenetischen Position der korrespondierenden Organismen, hochkonserviert. Auch der Vergleich der Primärstruktur der von *pmoCAB* abgeleiteten Aminosäure-Sequenzen mit den bekannten homologen Proteinen methanotropher Bakterien erbrachte keine Hinweise auf eine mögliche hohe Methan-Affinität der pMMO des USC α . Beispielsweise erfolgte ein Vergleich mit der unter 4.1 diskutierten Konsensussequenz aller bekannten pMMO-1 und AMO-Sequenzen (vgl. 2.4.1; Daten nicht gezeigt). Das Fehlen solcher Hinweise muß jedoch nicht zwangsläufig als Indiz dafür interpretiert werden, das die pMMO der Mitglieder des USC α keine hohe Affinität zu Methan aufweist. Erst mit der Aufklärung der dreidimensionalen Struktur der pMMO sowie der an der Reaktion direkt beteiligten Aminosäuren wird eine endgültige Diskussion der hier vorgestellten Daten möglich werden.

Phylogenetische Analyse der pMMO

Mit der Aufklärung der pMMO-Operons von *M. acidiphila* und USC α und der unter 2.1 beschriebenen und unter 4.1 diskutierten Analyse von *pmoCAB1* und *pmoCAB2* von *Methylocystis* sp. Stamm SC2 wurde die Anzahl vollständig sequenzierter pMMO-Operons mehr als verdoppelt. Deshalb konnte nun erstmals, unter Einbeziehung repräsentativer AMO-Operons nitrifizierender Bakterien eine auf den vollständigen Sequenzen der pMMO beruhende Phylogenie vorgestellt werden. Die der Berechnung zugrundeliegende Sequenzinformation wurde dabei im Vergleich zur partiellen *pmoA* (durchschnittlich ca. 500 bp analysiert) versechsfacht.

Die berechnete Phylogenie von PmoCAB ähnelt sehr stark der Phylogenie partieller *pmoA*/PmoA-Sequenzen. Lediglich die exakte Verzweigungsordnung der pMMO-2 von *Methylocystis* sp. Stamm SC2 und der USC α /*M. acidiphila*-Linie weichen, unabhängig von den verwendeten Algorithmen und Evolutionsmodellen,

von auf partiellen Sequenzen basierenden *pmoA*/PmoA-Phylogenien ab. Die Tatsache, daß pMMOs von USC α und *M. acidiphila* nicht monophyletisch mit konventionellen pMMOs der Typ II MB verzweigen, unterstreicht die Eigenständigkeit dieser erst kürzlich beschriebenen MB innerhalb der alphaproteobakteriellen Methanotrophen. Um zu demonstrieren, daß die Verzweigungsordnung nicht lediglich ein durch den verwendeten kleinen Datensatz hervorgerufenes Artefakt ist, wurden die Sequenzen der einzelnen Polypeptide unabhängig voneinander verrechnet. Das Verzweigungsmuster der Phylogenien partieller PmoA-Sequenzen wurde bei der Analyse des kleinen Datensatzes vollständiger PmoA-Sequenzen bestätigt. Die tiefe Verzweigung der pMMO von USC α und *M. acidiphila* wird somit nicht durch den eingeschränkten Datensatz hervorgerufen. Dagegen konnte sie hauptsächlich auf den Einfluß der PmoB-Sequenzen zurückgeführt werden. Der Nutzen des homologen Polypeptids AmoB für phylogenetische Berechnungen sowie potentielle Vorzüge gegenüber der Verwendung von AmoA-Sequenzen wurde kürzlich von Calvo und Garcia (2004) beschrieben. Die veränderte Verzweigungsordnung gewährt neue Einsichten in die Evolution der alphaproteobakteriellen MB sowie insbesondere der neuartigen Form der pMMO, pMMO-2.

Ein mögliches viertes Gen im pMMO-Operon

Bei der vergleichenden Analyse der pMMO-Operons zweier Typ II MB identifizierten Gilbert und Mitarbeiter (2000) unmittelbar stromabwärts des *pmoB*-Gens von *M. trichosporium* OB3b einen ORF ("*orfD*", partiell sequenziert), der hohe Sequenz-Ähnlichkeiten zu einem häufig stromabwärts des *amoB*-Gens nitrifizierender Bakterien gelegenen ORFs besitzt. Aufgrund der konservierten Lage in unmittelbarer Nähe der drei *amo*-Gene wurde dieser ORF als "*orf4*" annotiert. Stromabwärts des pMMO-Operons von *Methylocystis* sp. Stamm M konnten Gilbert und Mitarbeiter kein entsprechendes Gen nachweisen. Ebenso sind unterhalb der in 2.1 beschriebenen pMMO-Operons von *Methylocystis* sp. Stamm SC2 wie auch im analysierten Genom-Fragment von *M. acidiphila* keine vergleichbaren Gene vorhanden. Ein zu *orfD* bzw. *orf4* homologes Gen, ORF45, liegt jedoch stromabwärts des *pmoB*-Gens des USC α -Fragments. Interessanterweise weist ORF45 wesentlich höhere Sequenz-Ähnlichkeiten zu einem unterhalb der jeweiligen Kopie des *orf4* von *N. europaea* gelegenen Gen auf ("*orf5*"; NE2060 bzw. NE0941), als zu dem homologen Gen von

M. trichosporium. Im Genom von *N. europaea* sind zwei AMO-Operons mit entsprechenden flankierenden Genen vorhanden. Aufgrund der Ähnlichkeitswerte sowie des Verbreitungsmusters von ORF45-Homologen kann geschlossen werden, daß die Reihenfolge *pmoC*, *pmoA*, *pmoB*, *Orf4*, *Orf5* eine ursprüngliche Organisation darstellt, die schon vor der Entstehung paraloger Kopien des Genclusters durch Gen-Duplikation innerhalb eines Genoms vorhanden war. Die Tatsache, daß die homologen Gene verschiedener MB und Nitrifizierer höchste Sequenz-Ähnlichkeiten entweder zu *orf4* oder zu *orf5* aufweisen (Daten nicht gezeigt) deutet stark darauf hin, dass beide Kopien bereits vor der Aufspaltung in AMO-Gene und pMMO-Gene vorhanden waren und das in verschiedenen Genomen jeweils eine der Gen-Kopien verloren ging. Die hohe Sequenz-Konserviertheit sowohl innerhalb eines Genoms (NE2060 und NE0941 sind zu 99,6 % identisch, die abgeleiteten Proteine stimmen vollständig überein) wie auch zwischen verschiedenen Gattungen (Ne2060 und ORF45 sind nicht weniger ähnlich als beispielsweise die hoch-konservierten Gene *pmoCAB* und *amoCAB*) deutet auf einen starken selektiven Druck hin.

In bisherigen Arbeiten zur Operonstruktur von *pmoCAB* bzw. *amoCAB* (Gilbert et al., 2000; Klotz et al., 1998) wurde davon ausgegangen, daß das stromabwärts von *pmoB/amoB* gelegene Gen kein Bestandteil des jeweiligen Operons ist. Damit scheint auch der Befund in Einklang zu stehen, daß zwischen *pmoB* und ORF45 des USC α -Fragments ein putativer Terminator identifiziert werden konnte. Die hohe Übereinstimmung sowohl der Sequenzen von *orf4/orf5/ORF45* als auch insbesondere der Lage in direkter Nachbarschaft der *pmoCAB/-amoCAB*-Operons spricht allerdings stark für eine physiologische Funktion des Gens. Desweiteren wurde eine Ribosomen-Bindestelle (Shine-Dalgarno-Sequenz) direkt stromaufwärts sowie ein zweiter Terminator stromabwärts des ORF45 nachgewiesen. Diese Befunde sprechen ebenfalls für die Expression des ORF45. Andererseits konnte im Bereich der ca. 100 bp zwischen ORF45 und *pmoB* kein Promoter identifiziert werden. Eine mögliche Transkription von Orf45 muß somit durch eine der beiden nachfolgenden Hypothesen erklärt werden:

- i) ORF45 ist (fakultativer) Teil des *pmoCAB*-Operons und wird beispielsweise durch Antitermination reguliert.
- ii) ORF45 wird selbstständig, unabhängig von *pmoCAB* reguliert. Gegen diese Annahme spricht neben der konservierten Lage in unmittelbarer Nähe des

pMMO-Operons die Tatsache, daß kein Promoter unterhalb des *pmoB*-Gens identifiziert wurde, sodaß ein möglicher Promoter also innerhalb des *pmoCAB*-Operons lokalisiert sein müßte.

Das Fehlen der entsprechenden Gene unmittelbar stromabwärts der pMMO-Operons anderer MB muß nicht gleichbedeutend mit dem Fehlen im gesamten Genom sein und spricht deshalb nicht unbedingt gegen die genannten Hypothese. So liegt beispielsweise das in dieser Arbeit im Genom von *M. capsulatus* Bath, dem einzigen vollständig sequenzierten Methanotrophen-Genom, identifizierte Homologon zu ORF45 (MCA2130) auch nicht in unmittelbarer Nähe eines pMMO-Operons.

4.3 Die Nitrogenase-Gene methanotropher Bakterien

4.3.1 Einleitung

MB sind in terrestrischen und aquatischen Habitaten weit verbreitet und haben aufgrund ihrer Befähigung zur Fixierung molekularen Stickstoffs eine wichtige Funktion in Stickstoff-armen Habitaten. Aufgrund physiologischer und morphologischer Besonderheiten können die kürzlich beschriebenen methanotrophen Gattungen *Methylocapsa* und *Methylocella* nicht den klassischen Typ II MB zugeordnet werden (vgl. 1.2). Darüberhinaus gruppieren die Vertreter dieser Gattungen in 16S rDNA-basierten Phylogenien nicht direkt mit den ebenfalls alphaproteobakteriellen Typ II MB der *Methylosinus/Methylocystis*-Gruppe sondern mit acidophilen heterotrophen Bakterien der Gattung *Beijerinckia*. Auch aufgrund der Fähigkeit zur N₂-Fixierung unter atmosphärischen O₂-Konzentrationen erinnert *M. acidiphila* stärker an *Beijerinckia* spp. als an andere MB. Ziele der nachfolgend diskutierten Arbeiten waren i) die 16S-rDNA-unabhängige Untersuchung der phylogenetischen Beziehung dieser metabolisch unterschiedlichen Organismen-Gruppen, ii) die Zuordnung von in der Umwelt detektierten *nif*-Sequenzen zu methanotrophen bakterien zu ermöglichen und iii) die molekulargenetischen Grundlagen der N₂-Fixierung in MB vergleichend zu analysieren. Deshalb wurden partielle Sequenzen der für die Dinitrogenase-Reduktase und für die α -Untereinheit der Dinitrogenase kodierenden Gene *nifH* und *nifD* verschiedener Typ I und Typ II MB sowie aller bekannten Vertreter der Gattungen *Methylocapsa* und *Methylocella* erhoben, mit bekannten Sequenzen aus öffentlichen Datenbanken phylogenetisch verrechnet und im Kontext physiologischer Daten vergleichend diskutiert.

4.3.2 Vergleichende phylogenetische Analyse von NifH und NifD

In der letzten Dekade wurde eine Vielzahl kultivierungsunabhängiger *nifH*-basierter Studien diazotropher Mikroorganismen durchgeführt (Ueda et al., 1995; Zehr et al., 1995; Zehr et al., 1998; Rösch et al., 2002). Die Verwendung von *nifH* als funktioneller Genmarker wird durch die hohe Konserviertheit der *nifH*-Sequenzen

ermöglicht (Young, 1992; Martinez-Romero, 2000). Die Interpretation der Ergebnisse solcher Studien wird jedoch dadurch erschwert, daß nur relativ wenige *nif*-Gene kultivierter Organismen für die vergleichende Einordnung der Umweltsequenzen zur Verfügung stehen. Die im Zuge dieser Arbeit erhobenen Sequenzen sind die ersten aus MB isolierten *nifD*-Sequenzen. Darüberhinaus wurde die Anzahl öffentlich zugänglicher *nifH*-Sequenzen methanotropher Bakterien verdoppelt.

Aus DNA-Sequenzen abgeleitete Aminosäure-Sequenzen dienten als Basis für die Rekonstruktion der Phylogenie von NifH und NifD. Dazu wurden verschiedene Algorithmen und Korrekturfaktoren verwendet und die resultierenden Stammbäume anschließend manuell in einem Konsensusbaum verbunden, in den nur solche Verzweigungen übernommen wurden, die während aller Berechnungen unterstützt wurden (detailliertere Darstellung in 2.11.2). Die in dieser Arbeit erstellten Phylogenien von NifH, NifD und 16S rDNA weisen eine hohe Übereinstimmung auf. Allerdings gruppieren die Nif-Sequenzen einiger *Betaproteobacteria* (*Herbaspirillum seropedicae*, *Burkholderia fungorum*) innerhalb einer Gruppe von alphaproteobakteriellen Sequenzen. Dieses Muster wurde schon zuvor in NifH-basierten Phylogenien festgestellt und als starkes Indiz für die Übertragung von *nifH*-Genen mittels horizontalen Gentransfers (HGT) gewertet (Hurek et al., 1997). Die Kongruenz der NifD-basierten Stammbaum-Topologie unterstützt diese Interpretation und kann als Hinweis für eine Phylum-übergreifende Übertragung des gesamten *nifHDK*-Operons gewertet werden.

4.3.3 Ungewöhnliche Gruppierung von NifD und NifH von *M. capsulatus*

Ein weiteres Beispiel möglichen horizontalen Gentransfers alphaproteobakterieller *nif*-Gene stellt die Gruppierung der NifH- und NifD-Sequenzen des Typ X MB *M. capsulatus* (*Gammaproteobacteria*) innerhalb des von *Alphaproteobacteria* dominierten Sequenz-Clusters dar. Eine konvergente Evolution beider Polypeptide, NifD und NifH, erscheint als mögliche alternative Erklärung nicht plausibel. Eine vergleichende Analyse der relativen Verteilung aller im *nifHDK*-Operon von *M. capsulatus* für Aminosäuren kodierenden Trinukleotide („Codon-Usage“) gegen die Codon-Muster anderer, wahrscheinlich nicht durch horizontalen Gentransfer erworbener Operons (z.B. *pmoCAB*) konnte diese Annahme allerdings nicht bestätigen. Ein von der durchschnittlich in einem Genom beobachteten Codon-Usage

in kurzen Genomabschnitten abweichende Trinukleotidverteilung gilt als starker Hinweis auf HGT. Allerdings kann eine solche Abweichung nur nachgewiesen werden, wenn zwischen Donor und Rezipient ein statistisch signifikanter Unterschied der Codon-Usage besteht. Durch Vergleich der Codon-Usage verschiedener *Alphaproteobacteria* mit der durchschnittlichen Verteilung in *M. capsulatus* Bath konnte gezeigt werden, daß dies hier nicht der Fall ist (Daten nicht gezeigt). Die hohe Übereinstimmung der Codon-Usage dieser Organismen kann unter anderem auf den ähnlichen G+C-Gehalt ihrer Genome zurückgeführt werden, insofern G+C-Gehalt und Codon-Usage sich gegenseitig stark beeinflussen (Lipman und Wilbur 1983). Als Konsequenz scheint die Hypothese von der Aufnahme alphaproteobakterieller *nif*-Gene in das Genom von *M. capsulatus* durch Analyse der Codon-Usage prinzipiell weder falsifiziert noch bestätigt werden zu können. Hinzu kommt, daß der Nachweis lange zurückliegender Ereignisse horizontalen Gentransfers durch eine Selektionsbedingte Anpassung der übertragenen DNA an die Codon-Usage des Rezipienten erschwert wird: die Abweichung in der Nukleotid-Komposition sinkt mit der Zeit unter die Nachweisgrenze. Unbeachtet des evolutiven Ursprungs der *nif*-Gene im *M. capsulatus*-Genom stellt die atypische phylogenetische Position dieser Sequenzen neben dem für *Gammaproteobacteria* ungewöhnlichen G+C-Gehalt (*Methylococcus* spp.: 59-66%, Typ I MB: 49-60%, TypII MB 61-67%) das einzige molekulargenetische Beispiel für die Ähnlichkeit von *M. capsulatus* und Typ II MB dar (vgl. 1.2).

Die hier vorgestellten Phylogenien stehen in klarer Übereinstimmung mit physiologischen Studien. So wurde schon 1983 nachgewiesen, daß *Methylococcus* spp. sich durch die Fähigkeit, auf N-freiem Medium zu wachsen, von allen anderen gammaproteobakteriellen MB unterscheiden und in dieser Eigenschaft an alphaproteobakterielle Typ II MB erinnern, die bei O₂-Konzentrationen von bis zu 18% [v/v] zu N₂-Fixierung befähigt sind (Murrell and Dalton, 1983). Durch die gemeinsam mit diesen Phylogenien vorgestellten Wachstumsversuche konnte erstmals direkt gezeigt werden, daß auch Typ I MB in der Lage sind, molekularen Stickstoff zu fixieren und auf N-freiem Medium zu wachsen. Bisher war dies lediglich aufgrund indirekter Methoden (Nachweis von Acetylen-Reduktion, Existenz der *nif*-Gene) als phänotypisches Potential angenommen worden. Im Gegensatz zu *Methylococcus* und Typ II MB weisen Typ I MB eine deutlich höhere Sensitivität gegenüber Sauerstoff auf. Das Wachstum von Typ I MB konnte nur durch

regelmäßige Injektionen kleiner Mengen O_2 (O_2 -Konzentration $< 1 \%$ [v/v]) demonstriert werden, die eine für die Oxidation von Methan ausreichende O_2 -Versorgung gewährleisteten, ohne dabei die Nitrogenase vollständig zu inhibieren.

4.3.4 Phylogenetische Analyse von NifH und NifD acidophiler *Alphaproteobacteria*

Alle verrechneten NifD-Sequenzen acidophiler *Alphaproteobacteria* (Gattungen *Methylocapsa*, *Methylocella* und *Beijerinckia*) verzweigten innerhalb des von Sequenzen der *Alphaproteobacteria* dominierten NifD-Sequenzclusters. Wie auch in 16S rDNA-basierten Phylogenien bildeten die Vertreter dieser Gattungen eine monophyletische Gruppe. Die Sequenz-Identität der NifD-Sequenzen von *M. acidiphila* und *Beijerinckia* spp. lag jedoch mit Werten von mehr als 96% deutlich über den auf Basis der 16S-Phylogenie dieser Organismen zu erwartenden Werten. Die Ähnlichkeit der NifH-Sequenzen ist mit mehr als 98% Identität sogar noch höher. Zum Vergleich: die Identitäts-Werte sowohl der NifH- als auch der NifD-Sequenzen von *M. acidiphila* und *Methylocella* spp. liegen zwischen 90% und 92% (vgl. 3.3). Diese Befunde korrespondieren gut mit physiologischen Studien. Während *Beijerinckia* als „klassischer“ Diazotropher Stickstoff sogar bei atmosphärischen O_2 -Konzentrationen fixiert, werden die Nitrogenasen aller bekannten Typ I, Typ II und Typ X MB unter diesen Bedingungen inhibiert. Das einzige kultivierte MB, für welches dies nicht gilt, ist *M. acidiphila* (Dedysh et al., 2002). Dieses Bakterium zeigt, wie *Beijerinckia* spp., exponentielles Wachstum auf N-freiem Medium auch bei atmosphärischen O_2 -Konzentrationen und ist bis heute das einzige auf N-freiem Medium isolierte MB. *Beijerinckia* und *Methylocapsa* sind phylogenetisch nah verwandt. Die sehr hohe Übereinstimmung der Nif-Polypeptide von Vertretern beider Gattungen deutet darauf hin, daß der gemeinsame acidophile Vorfahre ebenfalls zur N_2 -Fixierung unter hohen O_2 -Konzentrationen befähigt war. Auf Ebene der DNA liegt die Ähnlichkeit sowohl der *nifH*- wie auch der *nifD*-Sequenzen bei 90%. Dies kann als Hinweis interpretiert werden, daß ein hoher Selektionsdruck auf die Polypeptid-Sequenzen einwirkte, sodaß in Relation zur Anzahl synonyme Mutationen nur wenige nichtsynonyme Mutationen akzeptiert wurden. Wäre die hohe Ähnlichkeit der Polypeptide auf horizontalen Gentransfer anstelle starken

konservierenden Selektionsdrucks zurückzuführen, sollten die Identitäts-Werte der entsprechenden Gene ebenfalls höher sein.

Die mit *M. acidiphila* und *Beijerinckia* spp nah verwandten Vertreter der Gattung *Methylocella* können, im Gegensatz zu *M. acidiphila*, neben anorganischen Verbindungen auch Methylamine und Aminosäuren als N-Quellen nutzen (Dunfield et al., 2003; Dedysh et al., 2004). Sie sind somit in den oft stickstoffarmen sauren Habitaten nicht obligat auf die N₂-Fixierung angewiesen. Dies könnte erklären, weshalb ein geringerer selektiver Druck auf Dinitrogenase und Dinitrogenase-Reduktase dieser MB wirkte, was in den hier diskutierten deutlich niedrigeren Identitätswerten auf Aminosäure-Ebene resultierte.

Zusammenfassend erlaubt die vergleichende Sequenzanalyse der *nifH*- und *nifD*-Gene Schlußfolgerungen bezüglich der Fähigkeit zur N₂-Fixierung methanotropher Organismen unabhängig von deren phylogenetischer Position. Dies konnte an den Beispielen *Methylococcus*, *Methylocapsa* und *Methylocella* verdeutlicht werden. Hohe Sequenz-Übereinstimmungen von NifD bzw. NifH verschiedener MB implizieren demnach vergleichbare physiologische Eigenschaften dieser Organismen. Die Erhebung eines Datensatzes von NifH- und NifD-Sequenzen methanotropher Bakterien erlaubte einen ersten Vergleich mit in Umwelt-Studien erhobenen Sequenzdaten. Die globale Rolle, die MB im N-Kreislauf spielen, wurde bisher noch nicht detaillierter untersucht. Hingegen konnte eine positive Korrelation zwischen der Befähigung zur N₂-Fixierung und der TCE-Abbaurate methanotropher Bakterien nachgewiesen werden (Chu and Alvarez-Cohen, 1999). Auch unter diesem Aspekt ist die Fähigkeit zur N₂-Fixierung deshalb von großer Bedeutung (Auman et al., 2001).

4.4 *In silico* tRFLP-Analyse in ARB mittels des "Tools"¹ tRFcut

4.4.1 Einleitung

Die tRFLP-Analyse (Yeager et al., 2004; Avrahami et al., 2002; Braker et al., 2001; Horz et al., 2001) ist ein molekularbiologisches Werkzeug zur Analyse der Diversität und Struktur mikrobieller Gemeinschaften (qualitative und semiquantitative Analyse). Der zeitlich aufwendigste Schritt ist die Zuordnung experimentell erzielter Restriktionsfragmente zu den zugehörigen Organismen. Das hierfür gängige, arbeitsintensive Verfahren ist das Erstellen einer Klon-Bibliothek aus PCR-Amplifikaten, die auch Grundlage der tRFLP-Analyse waren, sowie eine anschließende tRFLP-Analyse der Inserts vieler individueller Klone. Dieses Verfahren ist in der Regel für jedes untersuchte Habitat erneut durchzuführen. Nach Sequenzierung der einzelnen Klone kann eine Zuordnung der in der Umwelt-Analyse erhaltenen Restriktionsfragmente zu einzelnen Sequenzen, phylogenetischen Gruppen, oder, aufgrund der Mehrdeutigkeit einzelner Fragmentlängen, zu sogenannten Operativen Taxonomischen Einheiten (OTUs) erfolgen. Eine OTU beinhaltet dabei alle Spezies, für die identische Restriktionsfragment-Längen festgestellt wurden. Oft zeigt sich erst zu diesem Zeitpunkt, daß die verwendete Kombination von PCR-Primern und Restriktionsenzym keine phylogenetische Auflösung der Population des analysierten Habitats zuläßt, da beispielsweise abundante Fragmente phylogenetisch heterogenen OTUs entsprechen.

In-silico tRFLP-Analysen ermöglichen nicht bloß eine vereinfachte Zuordnung der erhaltenen Fragmente zu taxonomischen Gruppen, sondern auch die Verbesserung der phylogenetischen Auflösung eines Experimentes. Im Folgenden werden zunächst die generellen Möglichkeiten der *in silico* tRFLP-Analyse zur Lösung dieser Probleme diskutiert. Im Anschluß wird erläutert, welche Vorteile das im Rahmen dieser Arbeit entwickelte Tool ¹ tRFcut im Vergleich zu anderen tRFLP-Analyse-Programmen bietet. Darauf folgt eine kritische Betrachtung der *in silico* tRFLP-Analyse im Allgemeinen sowie von tRFcut im Besonderen. Die Diskussion

¹ Die in der Informatik gängige Bezeichnung "Tool" wird hier gewählt, da unter "Programm" eine eigenständig ausführbare Software verstanden wird, es sich bei tRFcut jedoch lediglich um eine Erweiterung ("Add-in") des Softwarepaketes ARB handelt.

endet mit einer kurzen Erörterung der Anwendung von tRFcut zur Analyse der methanotrophen Population eines pH-neutralen Waldbodens.

4.4.2 Zuordnung von Restriktionsfragmenten zu entsprechenden phylogenetischen Gruppen

Der Grundgedanke der *in silico* tRFLP-Analyse ist einfach: gebräuchliche Verfahren der Mustererkennung erlauben es, für eine gegebene DNA-Sequenz das unter Verwendung einer beliebigen Endonuklease zu erwartende terminale Restriktionsfragment zu berechnen. Die Analyse großer Gen-Datenbanken ermöglicht es innerhalb kürzester Zeit, die Restriktionsfragmente aller in der Datenbank vorhandenen Sequenzen (des entsprechenden Gentypus) zu berechnen. Das Zusammenfassen von Organismen mit sich entsprechenden Fragmentlängen zu OTUs ist anschließend ebenso leicht möglich, wie beispielsweise das gezielte Abfragen aller Sequenzen, die einem experimentell erzielten Fragment entsprechen. Wird dieses Verfahren auf große Datensätze (beispielsweise stark verbreitete Gene) angewendet, können allerdings Suchergebnisse erzielt werden, die einer gegebenen Fragmentlänge eine Vielzahl von Organismen zuordnen, die in dem untersuchten Habitat nicht vorhanden sind. Eine direkte Zuordnung ist dann ohne weitere Charakterisierung des Habitats nicht möglich. Allerdings ist eine Vielzahl von Fragestellungen denkbar, wo diese Kritik nicht oder nur bedingt greift, z.B.:

- **Wechsel von Primern oder Restriktionsenzym.** Wird während eines Experiments entschieden, eine Umstellung auf einen alternativen tRFLP-Assay vorzunehmen, beispielsweise wegen zu geringer (phylogenetischer) Auflösung des bisher verwendeten, können die bereits erhobenen habitatsspezifischen Sequenz-Daten *in silico* analysiert und anschließend auch für den neuen Assay genutzt werden. Das erneute Erstellen einer Klonbibliothek mit anschließender Bestimmung der Fragmentlängen einzelner Klone entfällt.
- **Verwendung gruppenspezifischer Primer.** Ist das der tRFLP zugrundeliegende PCR-System spezifisch für eine gut untersuchte, definierte Sequenzgruppe (phylogenetische Gruppe oder funktionelles Gen), kann die tRFLP-Analyse in gut charakterisierten, homogenen OTUs resultieren, die leicht *in silico* zu detektieren sind.

- **Überprüfung der Spezifität eines Fragmentes.** Mittels Klon-Bibliotheken wird immer nur eine Stichprobe der in einem Habitat vorhandenen Sequenzen analysiert. Um das Risiko abzuwägen, daß ein Fragment zu weiteren, lediglich zufallsbedingt nicht erzielten Sequenzen gehört, ist die Analyse einer großen Datenmenge notwendig. Dies kann durch eine *in silico* Analyse geschehen. Werden hierbei weitere, der OTU zuzuordnende Organismen gefunden, kann individuell die Wahrscheinlichkeit abgeschätzt werden, daß diese Organismen im gegebenen Habitat eine Rolle spielen. Daraufhin kann die Notwendigkeit des Erstellens einer größeren Klonbibliothek beurteilt werden.
- **Statistische Analysen von parallelen tRFLP-Analysen mit verschiedenen Enzymen.** Der in Zukunft vielleicht bedeutendste Aspekt der *in silico* tRFLP-Analyse zielt auf die komplette Substitution der arbeitsintensiven, Klon-basierten Zuordnungs-Methode ab: Kent und Mitarbeiter (2003) waren die ersten, die eine statistische Analysemethode beschrieben haben, die es erlaubt, die Resultate mehrerer, mittels verschiedener Enzyme durchgeführter tRFLP-Analysen eines Habitats *in silico* zu verrechnen. Das Programm ermittelt aus einer großen Sequenz-Menge diejenigen Sequenzen, deren *in silico* berechnete tRFs für *alle* verwendeten Enzyme mit einem der experimentell gewonnenen Fragmente übereinstimmen. Darüberhinaus soll es auch die quantitative Abschätzung der einzelnen Taxa anhand der in die Berechnung eingeflossenen experimentellen Daten („Peakhöhen“) erlauben. Die von Triplet veröffentlichte Software nimmt die *in silico* Analyse einer Datenbank nicht selbst vor, sondern erwartet Ergebnis-Tabellen, die auf anderem Wege (beispielsweise mittels tRFcut) generiert werden müssen.

4.4.3 Experiment-Design mittels *in silico* tRFLP-Analyse

Die Wahl des Restriktionsenzymys ist entscheidend für das Auflösungsvermögen einer tRFLP-Analyse. Wird beispielweise ein Enzym verwendet, das viele Sequenzen in einem hochkonservierten, nahe am markierten Primer gelegenen Bereich schneidet, wird die Mehrheit der analysierten Sequenzen eine identische Fragment-Länge aufweisen. Wird ein solches Problem erst bei der Auswertung der Experimente bzw. während der tRFLP-Analyse einer Klonbibliothek erkannt, ist die Wiederholung

vieler arbeitsintensiver Schritte unumgänglich. Eine geeignete Methode, hochauflösende Enzyme auszuwählen, ist die *in silico* tRFLP-Analyse. Dazu werden die Sequenzen der entsprechenden Gene aller in Frage kommenden Organismen mit einer Vielzahl von Enzymen (bzw. Kombinationen von Primern, Markierungsstelle und Enzymen) *in silico* analysiert. Können die Ergebnisse dieser Analyse strukturiert dargestellt werden (z.B. mittels Phylogenie-basiertem Clustering) ermöglichen sie eine schnelle Optimierung der tRFLP-Parameter. Da viele hundert Restriktionsenzyme zu Verfügung stehen, und darüberhinaus die Kombination von selten schneidenden Enzymen mit längeren Erkennungs-Sequenzen die Anzahl der möglichen Experimente noch potenziert, ist das Finden optimaler Parameter ohne computergestützte Berechnungen nahezu unmöglich.

4.4.4 tRFcut als Alternative zu anderen *in silico* tRFLP-Programmen

Da die individuelle Gegenüberstellung von tRFcut und anderen tRFLP-Programmen wie TAP-TRFLP (Marsh et al., 2000), torast (<http://www.torast.de>) und MiCA (<http://mica.ibest.uidaho.edu/>) den Rahmen dieser Diskussion sprengen würde, werden nachfolgend die wichtigsten Eigenschaften des Programms vorgestellt, die entweder gar nicht, zumindest aber in dieser Kombination bei keinem anderen Tool vorhanden sind.

- **Darstellung der Ergebnisse:** Die durch tRFcut berechneten tRFs werden als Einträge von Datenfeldern den analysierten Datensätzen („Spezies“) zugeordnet. Wie der Inhalt jedes anderen ARB-Datenfeldes können sie im Anschluß an die *in silico* Analyse manipuliert, sortiert, durchsucht oder exportiert werden. So ist es beispielweise möglich, alle Sequenzen, für die ein Restriktionsfragment bestimmter Länge berechnet wurde, zu selektieren und anschließend phylogenetisch zu analysieren. Die Darstellungen bereits vorhandener phylogenetischer Bäume können automatisch um berechnete tRF-Längen ergänzt werden. Eine mögliche Zuordnung einzelner Fragmentlängen zu phylogenetischen Gruppen ist somit sehr schnell möglich.
- **Arbeiten in gewohnter Umgebung / Analyse eigener Daten:** tRFcut wurde insbesondere entwickelt, da viele Wissenschaftler, die sich der tRFLP-Methode bedienen, auch mit dem Programm ARB vertraut sind. Sie verfügen

häufig über umfassende Bestände in ARB verwalteter, eigener wie öffentlicher Sequenzen. Eine beispielsweise über das Internet („remote“) durchgeführte *in silico* tRFLP-Analyse eigener Daten hat neben den weiter unten angesprochenen Punkten insbesondere den Nachteil, nur sehr umständlich handhabbar zu sein. Eigene Daten müssen aus der ARB-Datenbank exportiert und möglicherweise umformatiert, einem in der Regel unbekannten Service-Anbieter offenbart, und die Ergebnisse schließlich, häufig manuell, mit dem eigenen Datensatz verbunden werden. Diese Arbeitsschritte sind bei Erweiterung des Datensatzes zu wiederholen.

- **Beschränkungen öffentlicher Programme:** Die im Internet entweder mittels Browser („Remote Zugriff“) zugänglichen, oder frei kopierbaren („Installation“) Programme sind in ihrer Funktionalität häufig limitiert. Oft ist die Anzahl der für die Analyse nutzbaren Enzyme beschränkt, eigene (proprietäre) Daten können nicht in die Berechnung einbezogen werden oder nur die gängigsten Primer sind zulässig. Individuelle Fragestellungen können somit nicht beantwortet werden. Ein besonderes Problem entsteht, wenn funktionelle Gene (im Gegensatz zu 16S/18S rDNA-Sequenzen) analysiert werden sollen. Bislang war keine Software verfügbar, die solche Analysen ermöglichen würde. Alle diese Nachteile umgeht tRFcut. So kann der Nutzer beispielsweise unter allen nicht-redundant schneidenden Restriktionsenzymen auswählen, die zur Zeit in Rebase (Roberts et al., 2003), der umfassendsten Restriktionsenzym-Datenbank, vorhanden sind. Bei der Auswahl der Primer sind keinerlei Beschränkungen vorhanden, somit können auch selbst entwickelte Primersysteme in die Analyse einbezogen werden. Desweiteren ist tRFcut auf die Daten jeder in ARB verwalteten DNA (RNA)- Datenbank (vgl. auch 2.11.1) anwendbar.
- **Unabhängigkeit von der Primersequenz:** Es gehört zur gängigen Praxis, Primer-Sequenzen von neu erhobenen Sequenz-Daten vor deren Einreichung in öffentliche Datenbanken zu entfernen, da die Primersequenz nicht zu 100 % der Sequenz des analysierten Organismus entsprechen muß, bei der Sequenzierung aber immer die Sequenz des Primer-Oligonukleotids gelesen wird. Für die *in silico* tRFLP-Analyse resultiert daraus ein möglicherweise gravierendes Problem, da bei einer automatischen Primer-Erkennung durch die tRFLP-Software ein nicht geringer Teil der öffentlich verfügbaren

Sequenzdaten diskriminiert wird. tRFcut umgeht das Problem, in dem das dem verwendeten Datensatz zugrundeliegende Alignment für die Berechnung der Primer-Positionen herangezogen wird. Sequenzen, deren Primer-Bereich fehlt, die jedoch beginnend mit der dem Primer folgenden Base korrekt align sind, werden von tRFcut richtig erkannt. Ein Alignment der gesamten Sequenz ist nicht notwendig. So war es möglich, innerhalb weniger Stunden die partiellen Sequenzen von ca. 26.000 rRNA-Operons (partielle 28S rDNA, ITS1, 5S rDNA, ITS2, partielle 18S rDNA) von *Fungi* aus öffentlichen Datenbanken zu beziehen, daraus eine Datenbank in ARB zu erstellen, eine hochkonservierte, als Primerbindestelle für tRFLP-Analysen genutzte Region der 28S rDNA zu alignen und die tRF-Längen für alle ausreichend langen Sequenzen zu berechnen (Daten nicht gezeigt).

- **Ergebniskontrolle:** Sequenzen, die im analysierten Bereich nur unvollständig bekannt sind, werden nicht automatisch aus der Analyse ausgeschlossen, sondern dem Benutzer zusammen mit der Länge der Sequenz-Lücke angezeigt. Diese Sequenzen können generell von der weiteren Analyse ausgeschlossen oder die Daten individuell für jede Sequenz korrigiert werden. Fehlt beispielsweise ein in den Sequenzen aller nah verwandten Organismen hochkonservierter Bereich, kann eine berechtigte Annahme über die Basen-Abfolge der entsprechenden Sequenz (-Lücke) getroffen werden.

4.4.5 Allgemeine Probleme der *in silico* tRFLP-Analyse

Ein gravierendes, allen aktuellen tRFLP *in silico* Analysemethoden gemeinsames Problem ist die häufig beobachtete Abweichung experimentell erhobener Daten von den mittels Computern vorhergesagten Werten. Die experimentell bestimmten Längen der Restriktionsfragmente sind in Abhängigkeit von der verwendeten Analysesoftware, den genutzten Standards, den Analyseparametern sowie je nach individuell verwendeten Geräten und Chemikalien leichten Schwankungen unterworfen. Darüberhinaus können Abweichungen der experimentell bestimmten von den tatsächlichen Fragmentlängen aufgrund Sequenz-spezifischer Eigenschaften wie G+C-Gehalt, Länge der Fragmente bzw. möglicher Sekundärstrukturen auftreten (Daten nicht gezeigt). Beispielsweise sinkt der *relative* Längenunterschied zweier

nahezu gleich langer Fragmente mit deren *absoluter* Größe, weshalb lange Fragmente nur unzureichend aufgetrennt werden können. Es ist zu hoffen, daß eine mathematisch/statistische Analyse der Abweichungen von beobachteten und tatsächlichen Fragmentlängen eine Korrelation zu Sequenz-spezifischen Faktoren und somit eine mathematische Korrektur der experimentellen Daten ermöglichen wird. Beispielsweise konnten Kaplan und Kitts (2003), wenngleich nur auf Grundlage eines kleinen Datensatzes, eine positive Korrelation des „Shifts“ mit der Sequenzlänge, sowie eine negative Korrelation mit dem Puringehalt einer Sequenz nachweisen. Im Rahmen der vorliegenden Arbeit wurden verschiedene Analyse-Algorithmen sowie Sequenziergeräte getestet und solche Parameter ausgewählt, die eine möglichst geringe Verfälschung der Daten hervorriefen (Daten nicht gezeigt). Die beobachteten Abweichungen von ein bis zwei Basen können bei der Zuordnung der erzielten Restriktionsfragmente berücksichtigt werden.

4.4.6 tRFcut-spezifische Probleme

Ein Umschreiben des in Perl programmierten Quellcodes der ARB-Software hätte eine Anzahl gravierender Nachteile mit sich gebracht. Zum einen bestand die Gefahr, durch Seiteneffekte oder Programmierfehler unbeabsichtigt Fehler in anderen ARB-Algorithmen zu erzeugen. Darüberhinaus wäre jeder potentielle Nutzer des tRFcut-Tools gezwungen gewesen, die veränderte ARB-Software zu installieren (und dem Programmierer von tRFcut zu vertrauen).

Diese Probleme wurden umgangen, indem tRFcut als externes Shell-Skript realisiert wurde, das aus der ARB-Umgebung heraus aufgerufen wird. tRFcut erzeugt Nutzer-spezifische Befehle, die durch einen ARB-eigenen Interpreter (ARB command Interpreter, ACI) auf die Sequenz-Daten angewendet werden. Details sind in dem gemeinsam mit dem Programm tRFcut beziehbaren Handbuch (<http://www.uni-marburg.de/mpi/downloads/trfcut/trfcut.zip>) beschrieben. Nachteilig an dieser Lösung ist, daß die ARB-eigene Interpretersprache, in der diese automatisch erzeugten Nutzer-spezifischen Befehle formuliert sind, nur eine eingeschränkte Funktionalität bietet (geringe „Mächtigkeit“ der Sprache). Darüberhinaus wirken sich Fehler im ACI-Kode direkt auf tRFcut aus. Beispielsweise war es nicht möglich, Enzyme mit redundanten Erkennungssequenzen (‘N’) für die Analyse zu nutzen, da der ARB-Command-Interpreter das Benutzen von „Wildcards“ nicht in vollem Umfang

unterstützt. Desweiteren konnten aufgrund der externen Lösung wichtige ("mächtige") ARB-interne Werkzeuge, welche die Effektivität von tRFcut deutlich gesteigert hätten, (beispielsweise der PT-Server) ebenfalls nicht genutzt werden.

4.4.7 Anwendung von tRFcut zur Analyse methanotropher Bakterien in pH-neutralem Waldboden

Durch Anwendung molekularer Techniken wie der vergleichenden Sequenzanalyse des funktionellen Markergens *pmoA* konnte in einer Reihe von Studien die Abundanz der bislang nicht kultivierter MB-Gruppen USC α and USC γ in Böden nachgewiesen werden, die als Methan-Senken fungieren (Knief et al., 2003b; Jensen et al., 2000). Im Gegensatz dazu ist über die Verbreitung der als „Cluster I“ bezeichneten, kürzlich isolierten alphaproteobakteriellen MB (Pacheco-Oliver et al., 2002b) wenig bekannt (vgl. 1.2). Im Rahmen einer Studie waren *pmoA*-Sequenzdaten aus einem pH-neutralen, bei Göttingen gelegenen Boden erzielt worden, die dem Cluster I zugeordnet wurden (S. Kolb, pers. Mitteilung). Um die Abundanz der Gruppe in diesem Boden abzuschätzen, wurde eine tRFLP-Analyse durchgeführt. Für Datenauswertung und Experiment-Design wurde tRFcut in Kombination mit einer in ARB unterhaltenen Datenbank von rund 500 *pmoA*- und 900 *amoA*-Sequenzen genutzt. Zunächst wurde, ohne eine vorhergehende *in silico* Analyse durchzuführen, das bisher in allen publizierten *pmoA*-basierten tRFLP-Analysen verwendete Restriktionsenzym *MspI* in Kombination mit universellen Primern experimentell genutzt. Die Zuordnung der Fragmente mittels tRFcut resultierte in dem oben beschriebenen Problem mangelhafter phylogenetischer Auflösung der erzielten OTUs. Die hochabundanten Fragmente der Länge 79/80 Basen korrespondierten mit Vertretern verschiedener methanotropher Gruppen. Klonierung und Sequenzierung einer kleinen Stichprobe der PCR-Produkte ergab eine überraschend hohe Abundanz von Vertretern des Cluster I. Ein eindeutiger Nachweis der hohen Abundanz dieser Gruppe wäre allerdings nur durch die Analyse einer weitaus höheren Anzahl von Klonen möglich gewesen. Als Alternative für diese übliche Herangehensweise wurde eine *in silico* Analyse der Datenbank durchgeführt. Mittels tRFcut wurden für alle in der Datenbank enthaltenen Sequenzen *in silico* tRFLP-Analysen für verschiedene Enzyme durchgeführt und anschließend überprüft, ob die Auflösungsfähigkeit der

Enzyme eine Trennung der wichtigsten methanotrophen Gruppen, insbesondere der dem 79/80 Basen langen tRF zugeordneten Gruppen, ermöglichen würde. Dabei erwies sich das Enzym *HaeIII* als besonders gut geeignet. Aufgrund dieser Ergebnisse wurde eine weitere experimentelle tRFLP-Analyse der Umwelt-DNA mit *HaeIII* durchgeführt. Tatsächlich wurde das tRFLP-Muster von zwei tRFs mit Längen von 199 und 406 Basen dominiert, die aufgrund der *in silico* Analyse eindeutig dem Cluster I zugeordnet werden konnten. tRFcut berechnete Fragmente dieser Länge ausschließlich für *pmoA*-Sequenzen des Cluster I. Fragmente mit für andere methanotrophe Gruppen charakteristischen Längen wurden nicht oder nur mit sehr geringer Abundanz detektiert. Somit konnte eindeutig die hohe Abundanz von Vertretern des Cluster I in dem untersuchten Boden nachgewiesen werden. Eine experimentelle tRFLP-Analyse von Klonen, für die *in silico* tRFs der genannten Länge vorhergesagt wurden, bestätigte die Berechnungen (Abweichung +/- eine Base). Die Vorteile der *in silico* Analyse können daher wie folgt zusammengefaßt werden:

Die für diese Studie erstellte Klonbibliothek wurde hauptsächlich für die herkömmliche Herangehensweise ohne vorhergehende *in silico* Analyse benötigt. Bei Kenntnis genügend großer Sequenzmengen kann dieser Schritt umgangen werden. Beispielsweise kann der in dieser Arbeit entwickelte tRFLP-Assay nun zum spezifischen Nachweis von Cluster I in weiteren Habitaten genutzt werden, ohne daß das erneute Erstellen einer Klonbibliothek notwendig sein wird. Die anschließende Kontrolle der Ergebnisse der *in silico* Analyse diene lediglich der Überprüfung und Demonstration des neu erstellten Tools tRFcut. Insbesondere der Nachweis des nahezu ausschließlichen Vorkommens von Vertretern des Cluster I wäre ohne *in silico* Analyse nur durch erheblichen Sequenzier-Aufwand zu erzielen gewesen. Die hier vorgestellten Ergebnisse zeigen, daß unsere Kenntnis der methanotrophen *Bacteria* noch immer nicht alle umweltrelevanten Gruppen umfaßt. Die hohe Abundanz des Cluster I deutet auf eine besondere Anpassung dieser MB an das untersuchte Habitat hin. Worin diese Eigenschaften und Funktionen dieser Gruppe bestehen, wird Gegenstand weiterer Studien sein.

Literatur

Abe, T., Kanaya, S., Kinouchi, M., Ichiba, Y., Kozuki, T. und Ikemura, T. (2002). A novel bioinformatic strategy for unveiling hidden genome signatures of eukaryotes: self-organizing map of oligonucleotide frequency. *Genome* **13**, 12-20.

Abe, T., Kanaya, S., Kinouchi, M., Ichiba, Y., Kozuki, T. und Ikemura, T. (2003). Informatics for unveiling hidden genome signatures. *Genome Research*. 2003 Apr; **13**, 693-702.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. und Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-10.

Auman, A.J., Speake, C.C. und Lidstrom, M.E. (2001). *nifH* sequences and nitrogen fixation in type I and type II methanotrophs. *Applied and Environmental Microbiology* **67**, 4009-16.

Avrahami, S., Conrad, R. und Braker, G. (2002). Effect of soil ammonium concentration on N₂O release and on the community structure of ammonia oxidizers and denitrifiers. *Applied and Environmental Microbiology* **68**, 5685-92.

Banker, B.C., Kudze, H.K., Alford, D.P., DeLaune, R.D. und Lindau, C.W. (1995). Methane sources and sinks in paddy rice soil: relationship to emission. *Agriculture, Ecosystems and Environment* **53**, 243-51

Basu, P., Katterle, B., Andersson, K.K. und Dalton, H. (2003). The membrane-associated form of methane mono-oxygenase from *Methylococcus capsulatus* (Bath) is a copper/iron protein. *Biochemical Journal* **369**, 417-27.

Bender, M. und Conrad, R. (1992). Kinetics of CH₄ oxidation in oxic soils exposed to ambient air or high CH₄ mixing ratios. *FEMS Microbiology Ecology* **101**, 261-70.

Bender, M. und Conrad, R. (1993). Kinetics of CH₄ oxidation in oxic soils. *Chemosphere* **26**, 687-96.

Bender, M. und Conrad, R. (1995). Effect of CH₄ concentrations and soil conditions on the induction of CH₄ oxidation activity. *Soil Biology & Biochemistry* **27**, 1517-27.

Birren, B., Mancino, V. und Shizuya, H. (1999). Bacterial Artificial Chromosomes in Genome Analysis: A Laboratory Manual. Vol. 3, pp. 41-295, Cold Spring Harbor Press, Cold Spring Harbor, New York, USA.

Bodrossy, L., Holmes, E.M., Holmes, A.J., Kovacs, K.L. und Murrell, J.C. (1997). Analysis of 16S rRNA and methane monooxygenase gene sequences reveals a novel group of thermotolerant and thermophilic methanotrophs, *Methylocaldum* gen. nov. *Archives of Microbiology* **168**, 493-503.

Born, M., Dörr, H. und Levin, I. (1990). Methane consumption in aerated soils of the temperate zone. *Tellus* **42B**, 2-8.

- Bosse, U. und Frenzel, P.** (1997). Activity and distribution of methane-oxidizing bacteria in flooded rice soil microcosms and in rice plants (*Oryza sativa*). *Applied and Environmental Microbiology* **63**, 1199-207.
- Bowman, J.P., Skerratt, J.H., Nichols, P.D. und Sly, L.I.** (2001). Phospholipid fatty acid and liposaccharide fatty acid signature lipids in methane utilizing bacteria. *FEMS Microbiology Ecology* **85**, 15-22.
- Bowman, J.P., Jimenez, L., Rosario, I., Hazen, T.C. und Sayler, G.S.** (1993). Characterization of the methanotrophic bacterial community present in a trichloroethylene-contaminated subsurface groundwater site. *Applied and Environmental Microbiology* **59**, 2380-7.
- Bowman, J.P., McCammon, S.A. und Skerratt, J.H.** (1997). *Methylosphaera hansonii* gen. nov., sp. nov., a psychrophilic, group I methanotroph from Antarctic marine-salinity, meromictic lakes. *Microbiology* **143**, 1451-9.
- Braker, G., Ayala-del-Rio, H.L., Devol, A. H., Fesefeldt, A. und Tiedje, J. M.** (2001). Community structure of denitrifiers, Bacteria, and Archaea along redox gradients in Pacific Northwest marine sediments by terminal restriction fragment length polymorphism analysis of amplified nitrite reductase (*nirS*) and 16S rRNA genes. *Applied and Environmental Microbiology* **67**, 1893-901.
- Brendel, V. und Trifonov, E.N.** (1984). A Computer Algorithm for Testing Prokaryotic Terminators. *Nucleic Acids Research* **12**, 4411-27.
- Bull, I.D., Parekh, N.R., Hall, G.H., Ineson, P. und Evershed, R.P.** (2000). Detection and classification of atmospheric methane oxidizing bacteria in soil. *Nature* **405**, 175-8.
- Calvo, L. und Garcia-Gil, L.J.** (2004). Use of *amoB* as a new molecular marker for ammonia-oxidizing Bacteria. *Journal of Microbiological Methods* **57**, 69-78.
- Chan, S.I., Chen, K.H., Yu, S.S., Chen, C.L. und Kuo, S.S.** (2004). Toward delineating the structure and function of the particulate methane monooxygenase from methanotrophic bacteria. *Biochemistry* **43**, 4421-30.
- Chistoserdova, L., Gomelsky, L., Vorholt, J.A., Gomelsky, M., Tsygankov, Y.D. und Lidstrom, M.E.** (2000). Analysis of two formaldehyde oxidation pathways in *Methylobacillus flagellatus* KT, a ribulose monophosphate cycle methylotroph. *Microbiology* **146**, 233-8.
- Choi, D.W., Kunz, R.C., Boyd, E.S., Semrau, J.D., Antholine, W.E., Han, J.I., Zahn, J.A., Boyd, J.M., de la Mora, A.M. und DiSpirito, A.A.** (2003). The membrane-associated methane monooxygenase (pMMO) and pMMO-NADH: quinone oxidoreductase complex from *Methylococcus capsulatus* bath. *Journal of Bacteriology* **185**, 5755-64.
- Chu, K.H. und Alvarez-Cohen, L.** (1999). Evaluation of toxic effects of aeration and trichloroethylene oxidation on methanotrophic bacteria grown with different nitrogen sources. *Applied and Environmental Microbiology* **65**, 766-72.

- Cicerone, R.J. und Oremland, R.S.** (1988). Biogeochemical Aspects of Atmospheric Methane. *Global Biogeochemical Cycles* **2**, 299-327.
- Claros, M.G. und von Heijne, G.** (1994). TopPred II: An Improved Software For Membrane Protein Structure Predictions. *CABIOS* **10**, 685-6.
- Conrad, R. und Rothfuss, F.** (1991). Methane oxidation in the soil surface layer of a flooded rice field and the effect of ammonium. *Biology and Fertility of Soils* **12**, 28-32.
- Conrad, R.** (1996). Soil microorganisms as controllers of atmospheric trace gases (H₂, CO, CH₄, OCS, N₂O, and NO). *Microbiological Reviews* **60**, 609-40.
- Crutzen, P.J.** (1991). Methane's sinks and sources. *Nature*. **350**, 380-1.
- Cunnold, D.M., Steele, L.P., Fraser, P.J., Simmonds, P.G., Prinn, R.G., Weiss, R.F., Porter, L.W., O'Doherty, S., Langenfelds, R.L., Krummel, P.B., Wang, H.J., Emmons, L., Tie, X.X. und Dlugokencky, E.J.** (2002). In situ measurements of atmospheric methane at GAGE/AGAGE sites during 1985-2000 and resulting source inferences. *Journal of Geophysical Research-Atmospheres*. **107**, ACH 20.1-20.18.
- Deschavanne, P.J., Giron, A., Vilain, J., Fagot, G., Fertil, B.** (1999). Genomic signature: characterization and classification of species assessed by chaos game representation of sequences. *Molecular Biology and Evolution*. **16**, 1391-9.
- Dedysh, S.N., Khmelenina, V.N., Suzina, N.E., Trotsenko, Y.A., Semrau, J.D., Liesack, W. und Tiedje, J.M.** (2002). *Methylocapsa acidiphila* gen. nov., sp. nov., a novel methane-oxidizing and dinitrogen-fixing acidophilic bacterium from Sphagnum bog. *International Journal of Systematic and Evolutionary Microbiology* **52**, 251-61.
- Dedysh, S.N., Liesack, W., Khmelenina, V.N., Suzina, N.E., Trotsenko, Y.A., Semrau, J.D., Bares, A.M., Panikov, N.S. und Tiedje, J.M.** (2000). *Methylocella palustris* gen. nov., sp. nov., a new methane-oxidizing acidophilic bacterium from peat bogs, representing a novel subtype of serine-pathway methanotrophs. *International Journal of Systematic and Evolutionary Microbiology* **50**, 955-69.
- Dedysh, S.N., Panikov, N.S. und Tiedje, J.M.** (1998b). Acidophilic methanotrophic communities from Sphagnum peat bogs. *Applied and Environmental Microbiology* **64**, 922-9.
- Dedysh, S.N., Panikov, N.S., Liesack, W., Grosskopf, R., Zhou, J. und Tiedje, J.M.** (1998a). Isolation of acidophilic methane-oxidizing bacteria from northern peat wetlands. *Science* **282**, 281-4.
- Dedysh, S.N., Ricke, P. und Liesack, W.** (2004). NifH and NifD phylogenies: an evolutionary basis for understanding nitrogen fixation capabilities of methanotrophic bacteria. *Microbiology*. **150**, 1301-13
- Denier van der Gon, H.A.C. und Neue, H.U.** (1996). Oxidation of methane in the rhizosphere of rice plants. *Biology and Fertility of Soils* **22**, 359-66.

- Dörr, H., Katruff, L. und Levin, I.** (1993). Soil texture parameterization of the methane uptake in aerated soils. *Chemosphere* **26**, 697-713.
- Dunfield P. und Knowles R.** (1995). Kinetics of inhibition of methane oxidation by nitrate, nitrite, and ammonium in a humisol. *Applied and Environmental Microbiology* **61**, 3129-35.
- Dunfield, P.F., Khmelenina, V.N., Suzina, N.E. und Dedysh, S.N.** (2003). *Methylocella silvestris* sp. Nov. a novel methanotrophic bacterium isolated from an acidic forest cambisol. *International Journal of Systematic and Evolutionary Microbiology* **53**, 1231-9.
- Dunfield, P.F. und Conrad, R.** (2000). Starvation alters the apparent half-saturation constant for methane in the type II methanotroph *Methylocystis* strain LR1. *Applied and Environmental Microbiology* **66**, 4136-8.
- Dunfield, P.F., Liesack, W., Henckel, T., Knowles, R. und Conrad, R.** (1999). High-affinity methane oxidation by a soil enrichment culture containing a type II methanotroph. *Applied and Environmental Microbiology* **65**, 1009-14.
- Dunfield, P.F., Yimga, M.T., Dedysh, S.N., Berger, U., Liesack, W. und Heyer, J.** (2002). Isolation of a *Methylocystis* strain containing a novel *pmoA*-like gene. *FEMS Microbiology Ecology* **41**, 17-26.
- Edwards, S.V., Fertil, B., Giron, A. und Deschavanne, P.J.** (2002). A genomic schism in birds revealed by phylogenetic analysis of DNA strings. *Systematic Biology*. **51**, 599-613.
- Fjellbirkeland, A., Torsvik, V. und Ovreas, L.** (2001). Methanotrophic diversity in an agricultural soil as evaluated by denaturing gradient gel electrophoresis profiles of *pmoA*, *mxoF* and 16S rDNA sequences. *Antonie Van Leeuwenhoek*. **79**, 209-17.
- Fuse, H., Ohta, M., Takimura, O., Murakami, K., Inoue, H., Yamaoka, Y., Oclarit, J.M. und Omori, T.** (1998). Oxidation of trichloroethylene and dimethyl sulfide by a marine *Methylobacterium* strain containing soluble methane monooxygenase. *Bioscience, Biotechnology and Biochemistry* **62**, 1925-31.
- Georgiadis, M.M., Komiya, H., Chakrabarti, P., Woo, D., Kornuc, J.J. und Rees, D.C.** (1992). Crystallographic structure of the nitrogenase iron protein from *Azotobacter vinelandii*. *Science* **257**, 1653-9.
- Gilbert, B., McDonald, I.R., Finch, R., Stafford, G.P., Nielsen, A.K. und Murrell, J.C.** (2000). Molecular analysis of the *pmo* (particulate methane monooxygenase) operons from two type II methanotrophs. *Applied and Environmental Microbiology* **66**, 966-75.
- Gilbert, B. und Frenzel, P.** (1998). Rice roots and CH₄ oxidation: the activity of bacteria, their distribution and the microenvironment. *Soil Biology Biochemistry* **30**, 1903-16.
- Goodwin, P.M. und Anthony, C.** (1998). The biochemistry, physiology and genetics of PQQ and PQQ-containing enzymes. *Advances in Microbial Physiology* **40**, 1-80.

- Guirl, M.A. und Dooley, D.M.** (1999). Copper-containing oxidases. *Current Opinion in Chemical Biology* **3**, 138-44
- Hanson R.S. und Hanson, T.E.** (1996). Methanotrophic Bacteria. *Microbiological Reviews* **60**, 439-71.
- Harriss, R.C., Sebach, D.I. und Day, J.F.P.** (1982). Methane flux in the Great Dismal Swamp. *Nature* **297**, 673-4.
- Henckel, T., Jackel, U., Schnell, S. und Conrad, R.** (2000a). Molecular analyses of novel methanotrophic communities in forest soil that oxidize atmospheric methane. *Applied and Environmental Microbiology* **66**, 1801-8.
- Henckel, T., Roslev, P. und Conrad, R.** (2000b). Effects of O₂ and CH₄ on presence and activity of the indigenous methanotrophic community in rice field soil. *Environmental Microbiology* **2**, 666-79.
- Henikoff, S. und Henikoff, J.G.** (1992). Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences USA* **89**, 10915-9.
- Heyer, J., Berger, U., Hardt, M. und Dunfield, P.F.** *Methylohalobium crimeensis* gen. nov. sp. nov., a moderately halophilic methanotrophic bacterium isolated from hypersaline lakes of Crimea. Unpublished.
- Heyer, J., Galchenko, V.F. und Dunfield, P.F.** (2002). Molecular phylogeny of type II methane-oxidizing bacteria isolated from various environments. *Microbiology-Sgm* **148**, 2831-46.
- Holmes, A.J., Costello, A., Lidstrom, M.E. und Murrell, J.C.** (1995). Evidence that particulate methane monooxygenase and ammonia monooxygenase may be evolutionarily related. *FEMS Microbiology Letters* **132**, 203-8.
- Holmes, A.J., Roslev, P., McDonald, I.R., Iversen, N., Henriksen, K. und Murrell, J.C.** (1999). Characterization of methanotrophic bacterial populations in soils showing atmospheric methane uptake. *Applied and Environmental Microbiology* **65**, 3312-8.
- Hopp, T.P. und Woods, K.R.** (1981). Prediction of protein antigenic determinants from amino acid sequences. *Proceedings of the National Academy of Sciences USA* **78**, 3824-28.
- Horz, H.P., Yimga, M.T. und Liesack, W.** (2001). Detection of methanotroph diversity on roots of submerged rice plants by molecular retrieval of *pmoA*, *mmoX*, *mxaF*, and 16S rRNA and ribosomal DNA, including *pmoA*-based terminal restriction fragment length polymorphism profiling. *Applied and Environmental Microbiology* **67**, 4177-85.
- Hurek, T., Egner, T. und Reinhold-Hurek, B.** (1997). Divergence in nitrogenases of *Azoarcus* spp., Proteobacteria of the α subclass. *Journal of Bacteriology* **179**, 4172-8.

- Hütsch, W.B., Webster, C.P. und Powlson, D.S.** (1993). Long-term effect of nitrogen fertilization on methane oxidation in soil of the broadbalk wheat experiment. *Soil Biology Biochemistry* **25**, 1307-15.
- Jensen, S., Ovreas, L., Daae, F.L. und Torsvik, V.** (1998). Diversity in methane enrichments from agricultural soil revealed by DGGE separation of PCR amplified 16S rDNA fragments. *FEMS Microbiology Ecology* **26**, 17-26.
- Jensen, S., Holmes, A.J., Olsen, R.A. und Murrell, J.C.** (2000). Detection of methane oxidizing bacteria in forest soil by monoxygenase PCR amplification. *Microbial Ecology* **39**, 282-9.
- Jones, D. T., Taylor, W. R., und Thornton, J. M.** (1992). A new approach to protein fold recognition. *Nature*, **358**, 86-9.
- Jukes, T.H. und Cantor, C.R.** (1969). Evolution of Protein Molecules. In: *Mammalian Protein Metabolism*, (Munro, H.N., Ed.), Vol. III, 21-132, Academic Press, San Diego, California, USA.
- Kaplan, C.W. und Kitts, C.L.** (2003). Variation between observed and true Terminal Restriction Fragment length is dependent on true TRF length and purine content. *Journal of Microbiological Methods* **54**, 121–5.
- Karlin, S., Ladunga, I.** (1994). Comparisons of eukaryotic genomic sequences. *Proceedings of the National Academy of Science*. **91**, 12832–6.
- Karlin, S., Burge, C., Campbell, A.M.**(1992). Statistical analyses of counts and distributions of restriction sites in DNA sequences. *Nucleic Acids Research*. **20**, 1363–70.
- Karlin, S., Mrazek, J., Campbell, A.M.** (1997). Compositional biases of bacterial genomes and evolutionary implications. *Journal of Bacteriology*. **179**, 3899–913.
- Kent, A.D., Smith, D.J., Benson, B.J. und Triplett, E.W.** (2003). Web-based phylogenetic assignment tool for analysis of terminal restriction fragment length polymorphism profiles of microbial communities. *Applied and Environmental Microbiology* **69**, 6768-76. [<http://trflp.limnology.wisc.edu>].
- Kim, U.J., Birren, B.W., Slepak, T., Mancino, V., Boysen, C., Kang, H.L., Simon, M.I. und Shizuya, H.** (1996). Construction and characterization of a human bacterial artificial chromosome library. *Genomics*. **34**, 213-8.
- Kimura, M.** (1980). A Simple Method for Estimating Evolutionary Rates of Base Substitutions Through Comparative Studies of Nucleotide Sequences. *Journal of Molecular Evolution* **16**, 111-20.
- King, G.M.** (1993). Ecophysiological characteristics of obligate methanotrophic bacteria and methane oxidation in situ. In: Murrell, J.C. und Kelly, D.P., Eds. *Microbial growth on C₁-compounds*. Andover: Intercept. 303-13.

- Klotz, M.G. und Norton, J.M.** (1998). Multiple copies of ammonia monooxygenase (amo) operons have evolved under biased AT/GC mutational pressure in ammonia-oxidizing autotrophic bacteria. *FEMS Microbiology Letters* **168**, 303-11.
- Knief, C., Kolb, S., Bodelier, P.L.E., Lipski, A. und Dunfield, P.F.** Diversity and activity of methanotrophic bacteria in temporary flooded upland soils. In Vorbereitung.
- Knief, C.** (2004). Charakterisierung der methanotrophen lebensgemeinschaften in böden mit geringem methanangebot und der oligotrophen Adaption methanotropher Bakterien. Philipps-Universität Marburg. Dissertation.
- Knief, C., Altendorf, K. und Lipski, A.** (2003a). Linking autotrophic activity in environmental samples with specific bacterial taxa by detection of ¹³C-labelled fatty acids. *Environmental Microbiology* **5**, 1155-67.
- Knief, C., Lipski, A. und Dunfield, P.F.** (2003b). Diversity and activity of methanotrophic bacteria in different upland soils. *Applied and Environmental Microbiology* **69**, 6703-14.
- Knight, R.D., Freeland, S.J., Landweber, L.F.,** (2001). A simple model based on mutation and selection explains trends in codon and amino acid usage and GC composition within and across genomes. *Genome Biology*. **2**, research 0010.0–0010.13
- Kohonen, T.** (1990). The Self-Organizing Map. *Proceedings of the IEEE* **78**, 1464-80.
- Kolb, S.** (2004). Ein neuer quantitativer Ansatz zur Erfassung der Struktur und Aktivität methanotropher Gilden. Philipps-Universität Marburg. Dissertation.
- Kolb, S., Knief, C., Stubner, S. und Conrad, R.** (2003). Quantitative detection of methanotrophs in soil by novel pmoA-targeted real-time PCR assays. *Applied and Environmental Microbiology* **69**, 2423-9.
- Koschorrek, M. und Conrad, R.** (1993). Oxidation of atmospheric methane in soil: Measurements in the field, in soil cores and in soil samples. *Global. Biogeochemical. Cycles* **7**, 109-21.
- Krüger, M., Frenzel, P. und Conrad, R.** (2000). Microbial processes influencing methane emission from rice fields. *Global Change Biology* **6**, 1-15.
- Kyte J. und Doolittle R.F.** (1982). A Simple Method for Displaying the Hydropathic Character of a Protein. *Journal of Molecular Biology* **157**, 105-42.
- Lieberman, R.L., Shrestha, D.B., Doan, P.E., Hoffman, B.M., Stemmler, T.L. und Rosenzweig, A.C.** (2003). Purified particulate methane monooxygenase from *Methylococcus capsulatus* (Bath) is a dimer with both mononuclear copper and a copper-containing cluster. *Proceedings of the National Academy of Sciences USA* **100**, 3820-5.

- Lipman, D.J. und Wilbur, W.J.** (1983). Contextual constraints of synonymous codon choice. *Journal of Molecular Biology* **163**, 377–94.
- Lipscomb, J.D.** (1994). Biochemistry of the soluble methane monooxygenase. *Annual Review of Microbiology* **48**, 371-99.
- Lombardi, J.E., Epp, M.A. und Chanton, J.P.** (1997). Investigation of the methyl fluorid technique for determining rhizospheric methane oxidation. *Biogeochemistry* **36**, 153-72.
- Lontoh, S., Zahn, J.A., DiSpirito, A.A. und Semrau, J.D.** (2000). Identification of intermediates of in vivo trichloroethylene oxidation by the membrane-associated methane monooxygenase. *FEMS Microbiology Letters*. **186**, 109-13.
- Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Buchner, A., Lai, T., Steppi, S., Jobb, G., Forster, W., Brettske, I., Gerber, S., Ginhart, A.W., Gross, O., Grumann, S., Hermann, S., Jost, R., König, A., Liss, T., Lussmann, R., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A. und Schleifer, K.H.** (2004). ARB: a software environment for sequence data. *Nucleic Acids Research* **32**, 1363-71.
- Marsh, T.L., Saxman, P.R., Cole, J.R. und Tiedje, J.M.** (2000). Terminal restriction fragment length polymorphism analysis program, a web-based research tool for microbial community analysis. *Applied and Environmental Microbiology* **66**, 3616-20. [<http://rdp.cme.msu.edu>].
- Martinez-Romero, E.** (2000). Dinitrogen-fixing prokaryotes. In: *The Prokaryotes*. Edited by M. Dworkin. New York: Springer.
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. und Erlich, H.** (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Biotechnology*. **1992**, 17-27.
- Murrell, J.C., Gilbert, B. und McDonald, I.R.** (2000a). Molecular biology and regulation of methane monooxygenase. *Archives of Microbiology*. **173**, 325-32.
- Murrell, J.C. und Radajewski, S.** (2000). Cultivation-independent techniques for studying methanotroph ecology. *Research in Microbiology*. **151**, 807-14.
- Murrell, J.C. und Dalton, H.** (1983). Nitrogen fixation in obligate methanotrophs. *Journal of General Microbiology* **129**, 3481-6.
- Nanba, K. und King, G.M.** (2000). Response of atmospheric methane consumption by Maine forest soils to exogenous aluminum salts. *Applied and Environmental Microbiology* **66**, 3674-9.
- Needleman, S.B. und Wunsch, C.D.** (1970). A General Method Applicable to the Search for Similarities in the Amino Acid Sequence of Two Proteins. *Journal of Molecular Biology* **48**, 443-53.

- Nguyen, H.H.T., Nakagawa, K.H., Hedman, B., Elliott, S.J., Lidstrom, M.E., Hodgson, K.O. und Chan, S.I.** (1996). X-ray absorption and EPR studies on the copper ions associated with the particulate methane monooxygenase from *Methylococcus capsulatus* (Bath) - Cu(I) ions and their implications. *Journal of the American Chemical Society*. **118**, 12766-76.
- Nguyen, H.H., Elliott, S.J., Yip, J.H. und Chan, S. I.** (1998). The particulate methane monooxygenase from *Methylococcus capsulatus* (Bath) is a novel copper-containing three-subunit enzyme. Isolation and characterization. *Journal of Biological Chemistry* **273**, 7957-66.
- Norton, J.M., Low, J.M. und Klotz, M.G.** (1996). The gene encoding ammonia monooxygenase subunit A exists in three nearly identical copies in *Nitrosospira* sp. NpAV. *FEMS Microbiology Letters* **139**, 181-8.
- Ochsenreiter, T., Pfeifer, F. und Schleper, C.** (2002). Diversity of Archaea in hypersaline environments characterized by molecular-phylogenetic and cultivation studies. *Extremophiles*. **6**, 267-74.
- Oliver, J.L., Bernaola-Galvan, P., Guerrero-García, J., and Roman-Roldán, R.** (1993) Entropic profile of DNA sequences through chaos-gamederived images. *Journal of Theoretical Biology*. **160**, 457-70.
- Oldenhuis, R., Oedzes, J.Y., van der Waarde, J.J. und Janssen, D.B.** (1991). Kinetics of chlorinated hydrocarbon degradation by *Methylosinus trichosporium* OB3b and toxicity of trichloroethylene. *Applied and Environmental Microbiology* **57**, 7-14.
- Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. und de Jong, P.J.** (1998). An improved approach for construction of bacterial artificial chromosome libraries. *Genomics*. **52**, 1-8.
- Pacheco-Oliver, M., McDonald, I.R., Groleau, D., Murrell, J.C. und Miguez, C.B.** (2002). Detection of methanotrophs with highly divergent *pmoA* genes from Arctic soils. *FEMS Microbiology Letters* **209**, 313-9.
- Pau, R.N.** (1989). Nitrogenases without molybdenum. *Trends in Biochemical Sciences* **14**, 183-6.
- Paulus, V.** (2004) Untersuchungen zur wachstumsabhängigen Expression des *pmo1*- und *pmo2*-Operons in *Methylocystis* Stamm SC2. Philipps-Universität Marburg. Diplomarbeit
- Pearson, W.R.** (1990). Rapid and Sensitive Sequence Comparison with FASTP and FASTA Methods in *Enzymology* **183**, 63-98.
- Peterson, D.G., Tomkins, J.P., Frisch, D.A., Wing, R.A. und Paterson, A.H.** (2000). *Journal of Agricultural Genomics*, vol. 5. Construction of plant bacterial artificial chromosome (BAC) libraries: an illustrated guide. CABI Publishing, Cambridge, Mass.

- Rabus, R., Gade, D., Helbig, R., Bauer, M., Glockner, F.O., Kube, M., Schlesner, H., Reinhardt, R. und Amann, R.** (2002). Analysis of N-acetylglucosamine metabolism in the marine bacterium *Pirellula* sp. strain 1 by a proteomic approach. *Proteomics* **2**, 649-55.
- Reay, D.S., Radajewski, S., Murrell, J.C., McNamara, N. und Nedwell, D.B.** (2001). Effects of land-use on the activity and diversity of methane oxidizing bacteria in forest soils. *Soil Biology Biochemistry* **33**, 1613-23.
- Ren, T., Roy, R. und Knowles, R.** (2000). Production and consumption of nitric oxide by three methanotrophic bacteria. *Applied and Environmental Microbiology* **66**, 3891-7.
- Richardson, C.J., Tilton, D.L., Kadlec, J.A., Chamie, J.P.M. und Wentz, W.A.** (1978). Nutrient dynamics of northern wetland ecosystems. In *Freshwater wetlands – ecological processes and management potential* Edited by R.E. Good, D.F. Whigham & R.L. Simpson, New York: Academic Press, pp. 217-41.
- Ricke, P., Erkel, C., Kube, M., Reinhardt, R. und Liesack, W.** (2004). Comparative Analysis of the Conventional and Novel *pmo* (Particulate Methane Monooxygenase) Operons from *Methylocystis* Strain SC2. *Applied and Environmental Microbiology*. **70**, 3055–3063
- Ricke, P., Kolb, S. und Braker, G.** Application of a newly developed ARB-integrated *in silico* T-RFLP tool reveals the dominance of a novel *pmoA* cluster in a forest soil. *Applied and Environmental Microbiology*. Im Druck.
- Ricke, P., Erkel, C., Kube, M., Reinhardt, R. und Liesack, W.** USC α , a group of as-yet uncultured methanotrophs: First insights into the genomic organization substantiates an evolutionary affiliation to the alphaproteobacterial *Methylocapsa acidiphila*. In Vorbereitung
- Roberts, R.J., Vincze, T., Posfai, J. und Macelis, D.** (2003). REBASE: restriction enzymes and methyltransferases. *Nucleic Acids Research* **31**, 418-20.
- Rösch, C., Mergel, A. und Bothe, H.** (2002). Biodiversity of denitrifying and dinitrogen-fixing bacteria in an acid forest soil. *Applied and Environmental Microbiology* **68**, 3818-29.
- Roslev, P., Iversen, N. und Henriksen, K.** (1997). Oxidation and Assimilation of atmospheric methane by soil methane oxidizers. *Applied and Environmental Microbiology* **63**, 874-80.
- Roslev, P. und Iversen, N.** (1999a). Radioactive fingerprinting of microorganisms that oxidize atmospheric methane in different soils. *Applied and Environmental Microbiology* **65**, 4064-70.
- Saitou, N. und Nei, M.** (1987). The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. *Molecular Biology and Evolution* **4**, 406-25.

- Sandberg, R., Branden, C.I., Ernberg, I. und Coster, J.** (2003). Quantifying the species-specificity in genomic signatures, synonymous codon choice, amino acid usage and G+C content. *Gene* **311**, 35-42.
- Sandberg, R., Winberg, G., Branden, C.I., Kaske, A., Ernberg, I. und Coster, J.** (2001). Capturing whole-genome characteristics in short sequences using a naïve Bayesian classifier. *Genome Research* **11**, 1404-9.
- Sanger, F., Nicklen, S. und Coulson, A.R.** (1977). Sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences USA* **74**, 5463-7.
- Schnell, S. und King, G.M.** (1995). Stability of methane oxidation capacity to variations in methane and nutrient concentrations. *FEMS Microbiology Ecology* **17**, 285-94.
- Seiler, W., Conrad, R. und Scharffe, D.** (1984). Field studies of methane emission from termite nests into the atmosphere and measurements of methane uptake by tropical soils. *Journal of Atmospheric Chemistry* **1**, 171-86.
- Semrau, J.D., Chistoserdov, A., Lebron, J., Costello, A., Davagnino, J., Kenna, E., Holmes, A.J., Finch, R., Murrell, J.C. und Lidstrom, M.E.** (1995a). Particulate methane monooxygenase genes in methanotrophs. *Journal of Bacteriology* **177**, 3071-9.
- Semrau, J.D., Zolanz, D., Lidstrom, M.E. und Chan, S.I.** (1995b). The role of copper in the pMMO of *Methylococcus capsulatus* bath: a structural vs. catalytic function. *Journal of Inorganic Biochemistry* **58**, 235-44.
- Shigematsu, T., Hanada, S., Eguchi, M., Kamagata, Y., Kanagawa, T. und Kurane, R.** (1999). Soluble methane monooxygenase gene clusters from trichloroethylene-degrading *Methylobomonas* sp. strains and detection of methanotrophs during in situ bioremediation. *Applied and Environmental Microbiology* **65**, 5198-206.
- Shizuya, H., Birren, B., Kim, U.J., Mancino, V., Slepak, T., Tachiiri, Y. und Simon, M.** (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proceedings of the National Academy of Sciences USA* **89**, 8794-7.
- Smith, K.A., Dobbie, K.E., Ball, B.C., Bakken, L.R., Sitaula, B.K., Hansen, S., Klemmedtsson, L., Kasimir-Klemmedtsson, A., Degórska, A. und Orlanski P.** (2000). Oxidation of atmospheric methane in Northern European soils, comparison with other ecosystems, and uncertainties in the global terrestrial sink. *Global Change Biology* **6**, 791-803.
- Smith, T.F. und Waterman, M.S.** (1981). Comparison of Bio-Sequences. *Advances in Applied Mathematics* **2**, 482-9.
- Stainthorpe, A.C., Lees, V., Salmond, G.P., Dalton, H. und Murrell, J.C.** (1990). The methane monooxygenase gene cluster of *Methylococcus capsulatus* (Bath). *Gene* **91**, 27-34.

- Steudler, P.A., Bowden, R.D., Melillo, J.M. und Aber, J.D.** (1989). Influence of nitrogen fertilization on methane uptake in temperate forest soils. *Nature* **341**, 314-6.
- Stolyar, S., Costello, A.M., Peeples, T.L. und Lidstrom, M.E.** (1999). Role of multiple gene copies in particulate methane monooxygenase activity in the methane-oxidizing bacterium *Methylococcus capsulatus* Bath. *Microbiology* **145**, 1235-44.
- Stolyar, S., Franke, M. und Lidstrom, M.E.** (2001). Expression of individual copies of *Methylococcus capsulatus* bath particulate methane monooxygenase genes. *Journal of Bacteriology* **183**, 1810-2.
- Swofford D.L. und Olsen G.J.** (1990). Phylogeny reconstruction. In: *Molecular Systematics* (Hillis D.M. und Moritz C., eds.), pp. 411-501. Sinauer Associates, Sunderland.
- Takeguchi, M., Miyakawa, K. und Okura, I.** (1998). Properties of the membranes containing the particulate methane monooxygenase from *Methylosinus trichosporium* OB3b. *Biometals* **11**, 229-34.
- Tchawa Yimga, M., Dunfield, P.F., Ricke, P., Heyer, J. und Liesack, W.** (2003). Wide distribution of a novel pmoA-like gene copy among type II methanotrophs, and its expression in *Methylocystis* strain SC2. *Applied and Environmental Microbiology* **69**, 5593-602.
- Tchawa Yimga, M.** (2002). A novel type of *pmoA*: presence and distribution among methanotrophs - expression in *Methylocystis* strain SC2. Philipps-Universität Marburg. Dissertation.
- Thielemann, T.** (2000). Der Methanhaushalt über kohleführenden Sedimentbecken: Das Ruhrbecken und die Niederrheinische Bucht. Forschungszentrum Jülich. Dissertation.
- Tukhvatullin, I.A., Gvozdev, R.I. und Anderson, K.K.** (2001). Structural and functional model of methane hydroxylase of membrane-bound methane monooxygenase from *Methylococcus capsulatus* (Bath). *Russian Chemical Bulletin* **50**, 1867-76.
- Ueda, T., Suga, Y., Yahiro, N. und Matsuguchi, T.** (1995). Remarkable N₂-fixing bacterial diversity detected in rice roots by molecular evolutionary analysis of *nifH* gene sequences. *Journal of Bacteriology* **177**, 1414-7.
- Voet, D., Voet, J.G., Pratt, C.W.** (1998). *Fundamentals of Biochemistry*. John Wiley & Sons, New York, USA.
- von Heijne, G.** (1992). Membrane Protein Structure Prediction: Hydrophobicity Analysis and the 'Positive Inside' Rule. *Journal of Molecular Biology* **225**, 487-94.
- Vorholt J.A.** (2002). Cofactor-dependent pathways of formaldehyde oxidation in methylotrophic bacteria. *Archives of Microbiology*. **178**, 239-49.

- Wan, X.F., Xu, D., Kleinhofs, A. und Zhou, J.** (2004). Quantitative relationship between synonymous codon usage bias and GC composition across unicellular genomes. *BMC Evolutionary Biology* **4**, 19.
- Whalen, S.C. und Reeburgh, W.S.** (1990). Consumption of atmospheric methane by tundra soil. *Nature* **346**, 160-2.
- Whalen, S.C. und Reeburgh, W.S.** (1996). Moisture and temperature sensitivity of CH₄ oxidation in boreal soils. *Soil Biology Biochemistry* **28**, 1271-81.
- Whittenbury, R., Phillips, K.C. und Wilkinson J.F.** (1970). Enrichment, Isolation and some properties of methane-utilizing bacteria. *Journal of General Microbiology* **61**, 205-18.
- Wise, M.G., McArthur, J.V. und Shimkets, L.J.** (2001). *Methylosarcina fibrata* gen. nov., sp. nov. and *Methylosarcina quisquiliarum* sp.nov., novel type 1 methanotrophs. *International Journal of Systematic and Evolutionary Microbiology* **51**, 611-21.
- Wuebbles, D.J., Jain, A., Edmonds, J., Hayhoe, D. und Harvey, K.** (1999). Global change: state of the science. *Environmental Pollution* **100**, 57-86.
- Wuebbles, D.J. Hayhoe, K.** (2002). Atmospheric methane and global change. *Earth-Science Reviews* **57**, 177-210.
- Yavitt J.B., Downey D.M., Lancaster E. und Lang G.E.** (1990). Methane consumption in decomposition Sphagnum-derived peat. *Soil Biology and Biochemistry* **22**, 441-7.
- Yeager, C.M., Kornosky, J.L., Housman, D.C., Grote, E.E., Belnap, J. und Kuske, C.R.** (2004). Diazotrophic community structure and function in two successional stages of biological soil crusts from the Colorado Plateau and Chihuahuan Desert. *Applied and Environmental Microbiology* **70**, 973-83.
- Young, J.P.W.** (1992). Phylogenetic classification of nitrogen-fixing organisms,. In: *Biological nitrogen fixation*. Stacey, G., Burris, R.H. und Evans, H.J. (Eds.), Chapman & Hall, New York, N.Y. pp. 43-86.
- Yu, S.S., Chen, K.H., Tseng, M.Y., Wang, Y.S., Tseng, C.F., Chen, Y.J., Huang, D.S. und Chan, S.I.** (2003). Production of high-quality particulate methane monooxygenase in high yields from *Methylococcus capsulatus* (bath) with a hollow-fiber membrane bioreactor. *Journal of Bacteriology* **185**, 5915-24.
- Zahn, J.A., Bergmann, D.J., Boyd, J.M., Kunz, R.C. und DiSpirito, A.A.** (2001). Membrane-associated quinoprotein formaldehyde dehydrogenase from *Methylococcus capsulatus* Bath. *Journal of Bacteriology* **183**, 6832-40.
- Zahn, J.A. und DiSpirito, A.A.** (1996). Membrane-associated methane monooxygenase from *Methylococcus capsulatus* (Bath). *Journal of Bacteriology* **178**, 1018-29.

Zahn, J.A., Arciero, D.M., Hooper, A.B. und DiSpirito, A.A. (1996). Evidence for an iron center in the ammonia monooxygenase from *Nitrosomonas europaea*. *FEBS Letters* **397**, 35-8.

Zehr, J.P., Mellon, M., Braun, S., Litaker, W., Steppe, T. und Paerl, H.W. (1995). Diversity of heterotrophic nitrogen fixation genes in a marine cyanobacterial mat. *Applied and Environmental Microbiology* **61**, 2527-32.

Zehr, J.P., Mellon, M.T. und Zani, S. (1998). New nitrogen-fixing microorganisms detected in oligotrophic oceans by amplification of nitrogenase (*nifH*) genes. *Applied and Environmental Microbiology* **64**, 3444-50.

Zhou, J., Bruns, M.A. und Tiedje, J.M. (1996) DNA Recovery from Soils of Diverse Composition. *Applied and Environmental Microbiology*. **62**, 316–322

Lebenslauf

Persönliche Daten

Name:	Peter Ricke
Geburtstag:	30.11.1973
Geburtsort:	Neheim-Hüsten
Familienstand:	ledig

Schulausbildung

Aug. 1980 – Jun. 1984	Grundschule, Neheim
Aug. 1984 – Jun. 1990	Realschule, Arnsberg-Neheim
Aug. 1990 – Jun. 1993	Gymnasium, Höhere Berufsfachschule mit gymnasialer Oberstufe mit Schwerpunkt Wirtschaft und Verwaltung
Jul. 1993	Abitur

Ersatzdienst

Sep. 1993 – Sep. 1994	Pfarrei St. Michael, Neheim
-----------------------	-----------------------------

Studium

Okt. 1994 – Aug. 1996	Studium der Architektur, Universität Dortmund
Okt. 1996 – Aug. 2001	Studium der Biologie, Philipps-Universität Marburg
Okt. 1998	Vordiplom
Okt. 2000	Diplom-Hauptprüfung in den Fächern Mikrobiologie, Genetik und Biochemie
Nov. 2000 – Aug. 2001	Diplomarbeit am MPI für terrestrische Mikrobiologie in der Arbeitsgruppe von Dr. W. Liesack zum Thema „“
Juli 2001	Diplom
seit November 2001	Promotion am MPI für terrestrische Mikrobiologie in der Arbeitsgruppe von PD Dr. W. Liesack zum Thema „Molekularbiologische Untersuchungen zu Funktion und Phylogenie methanotropher Bakterien“

Dank

Ich bedanke mich bei Herrn Dr. Werner Liesack für die Überlassung des Themas und die häufigen, manchmal kontroversen, stets aber fruchtbaren Diskussionen über die methanotrophen Bakterien. Nicht zuletzt danke ich Ihm für die gewährten Freiräume und die vielen Möglichkeiten während der letzten vier Jahre eigene Ideen auszuprobieren.

Herrn Prof. Wolfgang Buckel danke ich für die Erstellung des Zweitgutachtens.

Frau Prof. Birgit Ziegenhagen und Herrn Prof. Uwe Maier danke ich für die Bereitschaft, mich zu prüfen.

Herrn Prof. Ralf Conrad danke ich für die Gelegenheit, in seiner Abteilung zu promovieren.

Bei Dr. Merlin Tschawa Yimga, Dr. Svetlana Dedysh, Christoph Erkel, Dr. Michael Kube, Dr. Richard Reinhardt, Dr. Gesche Braker und Dr. Steffen Kolb bedanke ich mich für die erfolgreiche Zusammenarbeit an den verschiedenen hier vorgestellten Themen.

Mein besonderer Dank gilt auch Sabrina Patzak und Julia Diehl für die große Hilfe bei der Verwaltung und dem Screening der Genom-Bibliotheken, insbesondere beim "Picken".

Natürlich denke ich in diesem Moment auch an alle Kollegen der Abteilung Biogeochemie, die heutigen und die gewesenen

Ganz besonders danke ich Christoph Erkel für die vielen kleinen und großen Dinge, die eine Freundschaft ausmachen.

Erklärung

Ich versichere, daß ich die Dissertation

Molekularbiologische Untersuchungen zu Funktion und Phylogenie methanotropher Bakterien

selbständig verfaßt, keine anderen als die angegebenen Hilfsmittel verwendet und sämtliche Stellen, die im Wortlaut oder dem Sinn nach anderen Werken entnommen sind, mit Quellenangaben kenntlich gemacht habe. Die Versicherung schließt Zeichnungen und Skizzen mit ein.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den 13. 11. 2004